PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C07H 21/00, 17/00, C12Q 1/68 — A1

(11) International Publication Number: **WO 00/47600**

(43) International Publication Date: 17 August 2000 (17.08.00)

(21) International Application Number: PCT/US00/03357

(22) International Filing Date: 10 February 2000 (10.02.00)

(30) Priority Data:

60/119,417	10 February 1999 (10.02.99)	US
60/136,185	27 May 1999 (27.05.99)	US

(71)(72) Applicants and Inventors: AL-SHEIKHLY, Mohamad [US/US]. 1013 Aster Boulevard, Rockville, MD 20850 (US). BENTLEY, William, E. [US/US]; P.O. Box 253, St. Michaels, MD 21601 (US). SILVERMAN, Joseph [US/US]; 8101 Connecticut Avenue, S-407, Chevy Chase, MD 20815 (US).

(74) Agents: ESMOND, Robert, W. et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100 New York Avenue, N.W., Washington, DC 20005-3934 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PHOTO-INDUCED NUCLEIC ACID HYBRIDIZATION

(57) Abstract

Photo-induced nucleic acid hybridization is achieved by exposure of a single-stranded nucleic acid molecules to ultraviolet (UV), visible (VIS) and near infrared (NIR) light. Specifically, irradiation increases the concentration of hydrogen bonded double-strand nucleic acid molecules as a result of complementary base pairing. Hybridization at pH 7.8 is most prevalent using UV (300 nm) irradiation, but is detectable even with NIR (920 nm) irradiation. The results offer promise of practical application in technologies related to genome arrays, northern and Southern blotting techniques, PCR, and hybrid nucleic acid-memory devices.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

3NSDOCID: <WO___0047600A1_I_>

Photo-Induced Nucleic Acid Hybridization

Field of the Invention

The invention is related to rapid nucleic acid hybridization by photo-induction at room temperature. Photo-induced nucleic acid hybridization can be used in the development of a nucleic acid memory chip for a hybrid system with massively parallel data searching capabilities wherein a directed light source induces hybridization reactions, and in nucleic acid amplification, e.g., via the polymerase chain reaction (PCR), and other techniques based on hybridization.

Background of the Invention

The fundamental chemistry of DNA is based on the double helix and the principal of complementarity. Hybridization is the hydrogen-bonding interaction between two DNA strands that obey the complementary rules. DNA strands can be very long, thread-like polynucleotides, made-up of a large number of deoxyribonucleotides containing purine and pyrimidine bases, which carry genetic information, and sugar and phosphate groups, which perform a structural role.

The variable part of DNA is its sequence of bases. DNA contains four kinds of bases, two purines and two pyrimidines. The two purines are adenine (A) and guanine (G). The two pyrimidines are thymine (T) and cytosine (C). Each strand of DNA is a covalently linked polymer, wherein each unit consists of a constant part (the sugar-phosphate "backbone") and one of either adenine, cytosine, guanine, or thymine. Each strand has a 3' and 5' end. When DNA forms a double-stranded helix, the strands must be anti-parallel (each end consists of a 3'-5' match) and complementary bases must align. Thus, a double-stranded helix will form through hydrogen bonding between the bases of two single strands. Hydrogen bonding occurs between A and T, and between G and C to produce complementary base pairs. This process is called annealing or hybridization. Ligation is the process in which the free 5' end of one strand reacts with the free 3' end of another strand to form a covalent bond to create a longer chain.

Although the structure of DNA helices is extremely rigid, it can be changed under certain conditions.

RNA, a single-stranded molecule, has a similar structure to DNA in terms of complementarity and polymeric structure. The variable part of RNA is also its sequence of bases. RNA contains four kinds of bases, two purines and two pyrimidines The two purines are adenine (A) and guanine (G). The two pyrimidines are uracil (U) and cytosine (C). Each strand of RNA is a covalently linked polymer, wherein each unit consists of a constant part (the sugar-phosphate "backbone") and one of either adenine, cytosine, guanine, or uracil. Each strand has a 3' and 5' end

The characteristics of nucleic acids are being manipulated in areas of technology such as DNA computing first described by Adleman, L.M., *Science 266:* 1021-1024 (1994), in an article entitled "Molecular computation of solutions to combinatorial problems." DNA computing addresses problems which require looking at several possible solutions (combinatorial problems) and solving those problems at the molecular level with nucleic acid molecules. There are many ongoing research efforts in DNA computing. *See, e.g.*, the extensive bibliography of molecular computing and splicing systems maintained by Pierluigi Frisco on the World Wide Web (*A Bibliography of Molecular Computation and Splicing Systems* (visited January 3, 2000) http://www.wi.leidenuniv.nl/~pier/dna.html). In addition, Baum has discussed a concept for an addressable memory system (Baum, E.B., *Science 268:* 583-585 (1995)). However, this memory system is in a flask and, as such, is not practical.

Among the ongoing efforts in DNA computing, the majority of the efforts are based upon computing in aqueous solutions. Among the principal activities in surface-based DNA computing are those being conducted at the University of Wisconsin, led by Max G. Lagally (Materials), Robert M. Corn (Chemistry) and Anne E. Condon (Computer Science). This Wisconsin group maintains a publications list of their own (Corn Group Publications (visited January 3, 2000) http://corninfo.chem.wisc.edu/publications.html). In a series of papers (Smith,

L.M., *et al.*, *J. Comput. Biol. 5:* 255-267 (1998); Frutos, A.G., *et al.*, *Nucl. Acids Res. 25:* 4748-4757 (1997); Cai, W., *et al.*, "The Power of Surface-Based Computation," *Proceedings of the First Annual International Conference on Computational Molecular Biology (RECOMB '97)*, Santa Fe, New Mexico (1997), Thiel, A.J., *Analytical Chem. 69:* 4948-4956 (1997)), the Wisconsin group has developed the rationale and foundation for surface-based DNA computing. Smith *et al.*, *J. Comput. Biol. 5:* 255-267 (1998), have indicated that "[c]omplex combinatorial mixtures of DNA molecules are immobilized on a surface and subsets are tagged and enzymatically modified in repeated cycles of the 'DNA computation.' When the computation is complete, the sequence of the DNA remaining is determined, yielding the computational result."

Smith *et al.* indicate that performing DNA operations on surfaces introduces some new issues of both a fundamental and practical nature. For example, one must take into account the fact that negligible diffusion rates for static molecules affect hybridization kinetics, that the net negative charge on or near the surface may repel negatively charged complements in the working medium, and that probes and complementary molecules may be inaccessible due to possible steric hindrance.

In addition, Smith *et al.* note that the use of all four bases (A, C, G, and T) to encode DNA is impractical because of the thermal stability problems occurring with a different G/C content. Encoding with two bases has the disadvantage of requiring longer oligomers to store data, but has the advantage of permitting the G/C content to be kept relatively constant.

Smith *et al.* also note that when synthesizing longer nucleotide sequences, the yield will be reduced exponentially with polymer length. For example, if each step (nucleotide attachment) has an efficiency of 99.5%, then synthesis of a 100-mer will produce only 0.995^{100} = 60% yield of the correct sequence. The remainder will be impurities and side products.

Frutos *et al.*, *Nucl. Acids Res. 25:* 4748-4757 (1997), have addressed the problem of finding the best sequences for storing data in oligonucleotides. They

considered the stability of oligonucleotides based on the percentage G/C content and optimal nucleotide order to minimize the number of partial matches during hybridization reactions.

Thiel *et al.*, *Analytical Chem. 69:* 4948-4956 (1997), were able to distinguish between single- and double-stranded DNA regions on a gold surface using the Surface Plasmon Resonance (SPR) imaging detection technique. The SPR technique requires no label on the DNA, but fluorescently-labeled targets were used for verification of this approach.

Another line of ongoing research is the development of self-assembled DNA monolayers on surfaces (Boland, T. and Ratner, B.D., *Langmuir 10*:3845-3852 (1994); Park, D., *et al.*, *Macromolecules 28:* 2595-2601 (1995); Pease, A.C., *et al.*, *Proc. Natl. Acad. Sci. USA 91*:5022-5026 (1994); McGall, G., *et al.*, *Proc. Natl. Acad. Sci. USA 93*:13555-13560 (1996)). Affymetrix, Inc. has developed a practical technique for "reading" the hybridized DNA pixels by stimulated fluorescence.

Clearly, hybridization plays in important role in new technologies, such as DNA computing as well as in technologies known in the art, such as PCR. Thus, a method of hybridization that would allow one to more efficiently and effectively carry out such technologies would be desirable. For example, in DNA computing, a compact DNA memory module would provide the levels of portability and durability that are needed for a practical hybrid system that are not now achievable.

Summary of the Invention

In the present invention, photo-induced hydrogen bonding between bases on single-stranded nucleic acid molecules is achieved not only by ultraviolet (UV), but also by visible (VIS) and near infrared (NIR) light. For example, while not wishing to be bound to any particular theory, two adjacent polynucleotides, which are otherwise stable in monomeric form in aqueous solutions, may be brought

BNSDOCID: <WO 0047600A1 | >

together by the exclusion of an excited water molecule (in the case of NIR) or by cation charge displacement (in the case of UV) so that hydrogen bonding occurs. This bonding can be detected, e.g., via gel electrophoresis and detectable markers such as fluorophore-tagged markers.

Thus, the invention relates to a method of inducing rapid hybridization at room temperature of one or more nucleic acid molecules comprising applying a photo-excitation source to the one or more nucleic acid molecules to induce hybridization.

In particular, the invention relates to a method of inducing nucleic acid hybridization comprising:

(a) immobilizing a monolayer of single-stranded nucleic acid molecules in a combinatorial array of pixels to a substrate;

(b) contacting the substrate with a mobile phase containing one or more target single-stranded nucleic acid molecules at least partially complementary to the immobilized nucleic acid molecules; and

(c) applying a photo-excitation source to at least one pixel position of the combinatorial array to induce hybridization between the immobilized nucleic acid molecules and the one or more target nucleic acid molecules.

The invention further relates to a method of inducing hybridization of two or more nucleic acid molecules which are at least partially complementary, comprising applying a photo-excitation source to the two or more nucleic acid molecules, whereby hybridization is induced.

The invention further relates to a method for amplifying at least one double-stranded nucleic acid molecule, comprising:

(a) providing a first and second primer, wherein the first primer is complementary to a sequence at or near the 3'-terminus of the first strand of the nucleic acid molecule to be amplified and the second primer is complementary to a sequence at or near the 3'-terminus of the second strand of the nucleic acid molecule to be amplified;

(b) mixing the first and second primers, and the nucleic acid molecule to be amplified, with one or more polypeptides having nucleic acid polymerase activity, e.g., one or more DNA polymerases, one or more RNA polymerases, one or more reverse transcriptases, etc., which may or may not be thermostable, thereby forming a mixture;

(c) illuminating the mixture with light under conditions favoring the photo-induced hybridization of the first primer to the first strand of the nucleic acid molecule, and the second primer to the second strand of the nucleic acid molecule;

(d) producing a third nucleic acid molecule complementary to the first strand and a fourth nucleic acid molecule complementary to the second strand;

(e) denaturing the first and third strands, and the second and fourth strands, and

(f) repeating steps (a) to (e) one or more times.

The invention further relates to a method of sequencing a nucleic acid molecule, comprising:

(a) mixing a primer with a first nucleic acid molecule to form a first mixture;

(b) illuminating the first mixture with a photo-excitation source under conditions favoring the photo-induced hybridization of the primer with the first nucleic acid molecule;

(c) mixing the first mixture with one or more deoxyribonucleoside triphosphates (dNTPs), one or more polypeptides having nucleic acid polymerase activity, e.g., one or more DNA polymerases, one or more RNA polymerases, one or more reverse transcriptases, etc., which may or may not be thermostable, and one or more terminator nucleotides, e.g., ddTTP, ddATP, ddGTP, ddITP or ddCTP, thereby forming a second mixture;

(d) incubating the second mixture under conditions sufficient to synthesize a random population of DNA molecules complementary to the first nucleic acid molecule, wherein the synthesized nucleic acid molecules are shorter

TRANSDOCID: <WO 0047600A1 I >

in length than the first nucleic acid molecule and wherein the synthesized nucleic acid molecules comprise a terminator nucleotide at their 3' termini; and

(e) separating the synthesized nucleic acid molecules by size so that at least a part of the nucleotide sequence of said first nucleic acid molecule can be determined.

The invention further relates to a nucleic acid memory device comprising a photo-excitation source, a substrate with nucleic acid molecules attached in a combinatorial array, a mobile phase containing nucleic acid target molecules at least partially complementary to the attached nucleic acid molecules, and a detector which detects hybridization within the combinatorial array.

Brief Description of the Figures

Fig. 1: A schematic diagram of a DNA memory module with an external photo-excitation source.

Fig. 2: A schematic diagram of a DNA memory module integrated with a semiconductor laser array.

Fig. 3A: A digital micromirror device that has two states - a "0" state wherein the mirror is tilted at +10 degrees and a "1" state wherein the mirror is tilted at -10 degrees. Fig. 3B: A schematic diagram of a DNA memory system using a DMD spatial light modulator.

Fig. 4: A DNA memory system employing a transmissive type SLM as the photo-excitation source.

Fig. 5: An illustration of storage and retrieval operations for DNA memory by sequencing.

Figs. 6A-6C: Photographs of hybridized DNA strands in aqueous solutions analyzed by gel electrophoresis. At each wavelength, an exposed sample, the same sample boiled, and the non-exposed control are shown.

Figs. 7A-7C: Photographs of hybridized DNA strands in aqueous solutions analyzed by gel electrophoresis.

ISDOCID: <WO___0047600A1_I_>

Fig. 8: A typical plot of the sum of the intensities of two bands plotted against the weight of an oligonucleotide.

Fig. 9A: A graph showing an increase in absorption (A_{260}) of DNA samples with respect to temperature. Fig. 9B: A photograph of melting depicted by gel electrophoresis.

Fig. 10: A graph showing the relative hybridization ratios of 37 μM oligonucleotide at pH 7.8 by photolysis in various regions: UV (300 nm), visible (450 nm) and NIR (920 nm) and the relative hybridization ratios of 37 μM oligonucleotide at various pH values. Note that the hybridization ratio (HR) was calculated as the intensity of light in gel (upper/lower) sample/ (upper/lower) unphotolyzed.

Fig. 11: A graph showing the effect on HR of the exposure time with 300 nm light on the hybridization yield of 37 µM at pH 7.8. Hybridization yield increases to a maximum value at 0.1 seconds. At higher exposure times (0.6 and 4.0 seconds), UV photolysis appears to induce scission of the hybridized molecules.

Fig. 12: A graph showing the effect on the HR of a DNA sample (pH= 7.8) photolyzed in UV, visible and IR regions.

Fig. 13: A graph showing the results of a DNA melting experiment. "A" represents absorbance at 260 nm which increases as temperature increases demonstrating that the double-stranded oligonucleotide is dissociated into two UV-light absorbing monomers. "B" represents the scanning density of the upper band in agarose gel which decreases demonstrating dissociation into two UV-absorbing monomers. A corresponding increase in the lower band was also observed (not shown).

Fig. 14: A diagram of the nucleotide sequence (SEQ ID NO:1) used in Example 1.

Detailed Description of the Preferred Embodiments

Typically, manipulation of DNA hybridization is made via changes in temperature. Higher temperatures cause melting of double-stranded DNA, while lower temperatures affect annealing, which are critical steps in PCR (U.S. Patent Nos. 4,683,195 and 4,683,202) and other technologies. Hybridization is also affected by buffer ionic strength and the use of detergents. The possibility that light can be used as a further control variable to enhance or otherwise tailor the annealing step has an enormous impact on the design and use of hybridization-based technologies such as techniques using genome arrays and northern and Southern blotting analytical techniques. Further, photo-induced hybridization substantially improves techniques based on enzyme extensions of nucleic acid hybridization such as PCR. Additionally, the realization of DNA-based computing (Adleman, L. M., *Science 266*:1021-1024 (1994); Frutos, A.G., *JACS 120*:10277-10282 (1998)) is greatly facilitated if a photochemical reaction can be found to induce hybridization of nucleic acid molecules efficiently since rapid photo-induced hybridization can provide a more realistic basis for nucleic acid-based information storage than slow enzyme-based chemistry. Thus, photonic interfaces (as shown in Fig. 2) would provide a degree of spatial control of nucleic acid hybridization which is unavailable by other means. In Fig. 2, the photonic interface is between the VCSEL array of the NIR emitting lasers and the target DNA gel.

Rapid hybridization at room temperature, in the manner described here, is useful for genomic chip-based technology, nucleic acid amplification and sequencing, and other laboratory techniques based on nucleic acid hybridization. The standard PCR technique, disclosed in U.S. Patent Nos. 4,683,195 and 4,683,202, is a technique that has been used for amplifying and sequencing nucleic acid molecules. In this technique, a sample containing the nucleic acid sequence to be amplified or sequenced (the "target sequence") is first heated to denature or separate the two strands of the nucleic acid. The sample is then cooled and mixed

with specific oligonucleotide primers which hybridize to the target sequence. Following this hybridization, for amplification of the target sequence, DNA polymerase in a buffered aqueous solution is added to the sample, along with a mixture of the dNTPs that are linked by the polymerase to the replicating nucleic acid strand. After allowing polymerization to proceed to completion, the products are again heat-denatured, subjected to another round of primer hybridization and polymerase replication, and this process repeated any number of times. Since each nucleic acid product of a given cycle of this process serves as a template for the production of two new nucleic acid molecules (one from each parent strand), the PCR process results in an exponential increase in the concentration of the target sequence. Thus, in a well-controlled, high-fidelity PCR process, as few as 20 cycles can result in an over one million-fold amplification of the target nucleic acid sequence.

By incorporating photo-induced hybridization in methods of nucleic acid amplification and sequencing, the annealing step performed after cooling at a reduced temperature may be eliminated. Instead, one may do the annealing and elongating at the same temperature. Additionally, one may tailor the specificity of various primers hybridized to the nucleic acid based on light excitation. That is, the incident light affords another control variable for fine tuning the amplification and sequencing reactions. This is an enormous advantage over the current state-of-the-art.

Other methods of nucleic acid amplification and sequencing may analogously employ the photo-induced hybridization methods of the invention. Examples of such amplification techniques include, but are not limited to, Strand Displacement Amplification (SDA; U.S. Patent No. 5,455,166; EP 0,684,315), and Nucleic Acid Sequence-Based Amplification (NASBA; U.S. Patent No. 5,409,818; EP 0 329 822). Examples of sequencing techniques include, but are not limited to, PCR-based nucleic acid fingerprinting techniques such as Random Amplified Polymorphic DNA (RAPD) analysis (Williams, J.G.K., *et al.*, *Nucl. Acids Res. 18(22):* 6531-6535 (1990)), Arbitrarily Primed PCR (AP-PCR; Welsh,

3NSDOCID: <WO__0047600A1_I_>

J., and McClelland, M., *Nucl. Acids Res. 18(24):* 7213-7218 (1990)), DNA Amplification Fingerprinting (DAF: Caetano-Anollés *et al.*, *Bio/Technology 9:* 553-557 (1991)), microsatellite PCR or Directed Amplification of Minisatellite-region DNA (DAMD; Heath, D.D., *et al.*, *Nucl. Acids Res. 21(24):* 5782-5785 (1993)), and Amplification Fragment Length Polymorphism (AFLP) analysis (EP 0 534 858, Vos, P., *et al.*, *Nucl. Acids Res. 23(21):* 4407-4414 (1995); Lin, J.J., and Kuo, J., *FOCUS 17(2):* 66-70 (1995)). In particular, the methods of the present invention will be useful in the fields of medical therapeutics and diagnostics, forensics, and agricultural and other biological sciences, in any procedure using hybridization of nucleic acid molecules.

Thus, in one embodiment, the invention relates to a method for amplifying at least one double-stranded nucleic acid molecule, comprising:

(a) providing a first and second primer, wherein the first primer is complementary to a sequence at or near the 3'-terminus of the first strand of the nucleic acid molecule to be amplified and the second primer is complementary to a sequence at or near the 3'-terminus of the second strand of the nucleic acid molecule to be amplified,

(b) mixing the first and second primers, and the nucleic acid molecule to be amplified, with one or more polypeptides having nucleic acid polymerase activity, e.g., one or more DNA polymerases, one or more RNA polymerases, one or more reverse transcriptases, etc., which may or may not be thermostable, thereby forming a mixture;

(c) illuminating the mixture with light under conditions favoring the photo-induced hybridization of the first primer to the first strand of the nucleic acid molecule, and the second primer to the second strand of the nucleic acid molecule;

(d) producing a third nucleic acid molecule complementary to the first strand and a fourth nucleic acid molecule complementary to the second strand;

(e) denaturing the first and third strands, and the second and fourth strands; and

(f) repeating steps (a) to (e) one or more times.

In preferred embodiments, the denaturing and hybridizing steps are performed at temperatures below about 90°C, below about 85°C, below about 80°C, below about 75°C, below about 70°C, below about 65°C, below about 60°C, below about 55°C, below about 50°C, below about 45°C, or below about 40°C, or at about 20°C to about 45°C.

In a related embodiment, the invention also provides a method of sequencing a nucleic acid molecule, comprising:

(a) mixing a primer with a first nucleic acid molecule to form a first mixture;

(b) illuminating the first mixture with a photo-excitation source under conditions favoring the photo-induced hybridization of the primer with the first nucleic acid molecule;

(c) mixing the first mixture with one or more deoxyribonucleoside triphosphates (dNTPs), one or more polypeptides having nucleic acid polymerase activity, e.g., one or more DNA polymerases, one or more RNA polymerases, one or more reverse transcriptases, etc., which may or may not be thermostable, and one or more terminator nucleotides, e.g., ddTTP, ddATP, ddGTP, ddITP or ddCTP, thereby forming a second mixture;

(d) incubating the second mixture under conditions sufficient to synthesize a random population of DNA molecules complementary to the first nucleic acid molecule, wherein the synthesized nucleic acid molecules are shorter in length than the first nucleic acid molecule and wherein the synthesized nucleic acid molecules comprise a terminator nucleotide at their 3' termini; and

(e) separating the synthesized nucleic acid molecules by size so that at least a part of the nucleotide sequence of said first nucleic acid molecule can be determined.

In a related aspect, the primer, first nucleic acid molecule, one or more dNTPs, one or more polypeptides having nucleic acid polymerase activity, and one or more terminator nucleotides may be mixed together prior to illuminating the

BNSDOCID: <WO__0047600A1_I_>

mixture with the photo-excitation source under conditions favoring the photo-induced hybridization of the primer with the first nucleic acid molecule, such that hybridization and synthesis of the random population of nucleic acid molecules may occur contemporaneously.

In addition to nucleic acid amplification and sequencing and other laboratory techniques, photo-induced hybridization can be used to obtain a high-density hybrid nucleic acid memory and massively parallel data searching system in an integrated, compact package, nucleic acid molecules, vectors containing such molecules, host cells containing such molecules, etc. Massively parallel searching is enabled by simultaneous hybridization reactions of all nucleic acid complementary pairs as described below.

Thus, another aspect of the invention relates to a method of inducing nucleic acid hybridization comprising:

(a) immobilizing a monolayer of single-stranded nucleic acid molecules in a combinatorial array of pixels to a substrate;

(b) contacting the substrate with a mobile phase containing one or more target single-stranded nucleic acid molecules at least partially complementary to the immobilized nucleic acid molecules; and

(c) applying a photo-excitation source to at least one pixel of the combinatorial array to induce hybridization between the immobilized nucleic acid molecules and the one or more target single-stranded nucleic acid molecules.

Another aspect of the invention relates to a method of inducing nucleic acid hybridization comprising:

(a) obtaining a substrate comprising a monolayer of single-stranded nucleic acid molecules in a combinatorial array of pixels immobilized thereon and a mobile phase containing one or more target single-stranded nucleic acid molecules at least partially complementary to the immobilized nucleic acid molecules; and

(b) applying a photo-excitation source to at least one pixel of the combinatorial array to induce hybridization between the immobilized nucleic acid molecules and the one or more target single-stranded nucleic acid molecules.

Photo-induced hybridization, according to this aspect of the invention, involves certain components in contact with, and layered upon, a semiconductor or other substrate. The components may include a photo-excitation source and its associated opto-electronics, a self-assembly of a monolayer of nucleic acid molecules immobilized on the substrate surface, and a mobile phase, e.g., a buffered salt solution, a gel matrix, etc., carrying solubilized nucleic acid at least partially complementary to the immobilized nucleic acid molecules. The nucleic acid molecules may be chemically affixed directly to the photo-excitation array or to an intermediate highly transmittive insulating layer such as fused silica.

The photo-excitation source may be internal or external to the substrate. As an example of an internal source, the substrate may comprise one more light-producing or emitting components that can be induced, e.g., electronically, electrically or physically, to produce light at a localized site, thereby hybridizing the nucleic acid molecules immobilized at that site. As an example of an external source, the substrate containing the immobilized nucleic acid molecules may be illuminated with a light source, thereby hybridizing the nucleic acid molecules. Detailed examples of internal and external light sources are provided below.

In one embodiment of the nucleic acid memory and searching system, the system is a nucleic acid memory device in the form of a computer chip interfaced with an electronic computer. Such chips, containing a potentially enormous database of hybridized nucleic acid molecules, with their near real-time massive parallel searching capabilities, can be applied to areas of target identification, parts location, and personnel identification, e.g., fingerprint and other bio-source identification, multi-level security, low power robotics with persistent memory (for powering down), "smart dust," i.e., memory units for nanoscale sensor arrays, "smart drugs," i.e., drugs that are able to respond to "invisible" conditions, among

INSDOCID: <WO___0047600A1_I_>

others. Additionally, photo-induced hybridization could be used to as a diagnostic technique in drug design, among other applications.

Photo-Induced Hybridization Device

A device according to the present invention may be used as a high-density hybrid nucleic acid memory and massively parallel data search system. The compact device may be assembled in a layered sandwich configuration comprising, in any order: a photo-excitation source (which is also possibly the substrate or is external to the device), a substrate with nucleic acid molecules immobilized thereon in a combinatorial array (pixels), a mobile phase containing target nucleic acid molecules complementary to the immobilized nucleic acid molecules, and a detector which senses hybridization at each pixel once the hybridization process is complete. The location of the detector may vary as shown in Figs. 1 and 2 which illustrate particular embodiments of the device. An electronic computer may be connected to the device to facilitate the data searching capability.

Substrate Layer and Internal Photo-Excitation Source

In particular, the substrate layer comprises well-defined, self-assembled nucleic acid molecules such as a monolayer of identical single-stranded nucleic acid molecules in a combinatorial array. This means that distinct pixel areas, i.e., areas containing particular nucleic acid molecules, on the substrate contain different nucleic acid sequences and that each pixel can be addressed individually. The self-assembly of the monolayer of nucleic acid molecules on the substrate surface must be accomplished with sufficient surface concentration and stability. This can be accomplished by placing a minimum of about 10^5 molecules/pixel (based on currently available pixel size well-known to one of ordinary skill in the art) and conducting the hybridization at room temperature.

Cai *et al.*, "The Power of Surface-Based Computation," *Proceedings of the First Annual International Conference on Computational Molecular Biology (RECOMB '97)*, Santa Fe, New Mexico (1997), note that good attachment chemistry ensures that the properly prepared nucleic acid molecules are immobilized to the substrate surface at a high density, and that other nucleic acid molecules exposed to the substrate surface later, for example, during hybridization, do not bond non-specifically to the substrate surface. Preliminary experiments on glass surfaces indicate that different chemistry can be used to attach either the 3' or 5' end of an oligonucleotide to the surface. However, one problem with glass is the non-specific binding of DNA molecules to the surface, i.e., the molecules bind to the surface itself rather than to their complement. Thus, thermally grown oxides on silicon wafers or alkanethiol self-assembled monolayers on gold surfaces may be better alternatives.

The substrate to which the monolayer is attached may be a photo-excitation source such as an integrated monolithic array of opto-electronic devices developed to assure a collimated monochromatic light source for transmission near the optimal wavelength which effectively induces nucleic acid hybridization, i.e., a wavelength between about 350 nm and about 1200 nm. The monochromatic light is the source of energy for the hybridization reactions between the nucleic acid molecules and the surrounding mobile phase medium. Preferred substrates include semiconductors, i.e., vertical cavity surface emitting laser (VCSEL) arrays. More preferably, the photo-excitation source includes semiconductor wafers such as GaAs or GaN. Compound semiconductors, such as GaAs, an excellent source of NIR light (850-950 nm), provide the activation energy for the hybridization reactions between the immobilized nucleic acid molecules and the surrounding target nucleic acid molecules. Towe *et al.* (Ramos, P.A. and Towe, E., *Appl. Phys. Lett. 69*:3321-3323 (1996); Towe, E. and Ramos, P.A., "Active (In, Ga) As/GaAs Blue-Green Light-Emitters," SPIE Photonics West, Optical Interconnects Symposium, San Jose, CA, Feb. 22, 1997; Ramos, P.A. and Towe, E., *Appl. Phys. Lett. 68:* 1754-1756 (1996); Ramos, P.A. and Towe, E., *Optics*

BNSDOCID: <WO__0047600A1_I_>

Comm. 132:121-127 (1996)) have shown that 112 GaAs wafers with [112]-oriented (In,Ga)As/GaAs epitaxial layers can produce coherent blue radiation (the frequency of the second harmonic). The semiconductor may also be used to detect differences in the electrical permittivity of the hybridized and unaltered pixels as described below.

External Photo-Excitation Source

As noted above, the energy required to induce hybridization may alternatively be provided by an external light source, such as a laser. It is necessary to determine the intensity and duration of the laser sufficient to hybridize enough immobilized molecules to be detectable. In order to identify the optimum wavelength, tunable lasers covering from about 260 nm in the UV range to about 10 nm in the NIR range, for example, may be used. An exemplary tunable laser can produce light pulses as short as 100 femtoseconds with peak power as high as several terawatts. Also, continuous sources of light with wavelengths in the optimum region may be used in PCR and blotting applications.

To efficiently deliver the light source for photo-induced hybridization, spatial light modulation (SLM) techniques may be applied to construct a programmable pattern generator that can be electronically programmed to illuminate specific areas on the surface so as to activate the hybridization process. Within the photo-excitation system, the SLM device controls the external light to address different pixels on the immobilized nucleic acid array. The SLM device can turn the light on and off at each pixel independently. There are two types of SLM devices: reflective type and transmissive type. Each type of SLM device has its own advantages and selection depends on the working wavelength. In either case, the selected SLM device may be controlled by a computer that also reads the data collected from the nucleic acid memory system.

One particular reflective type SLM device that may be used to photo-induce nucleic acid hybridization is the digital micro-mirror device (DMD)

NSDOCID: <WO___0047600A1_I_>

developed by Texas Instruments. The DMD is a reflective type device containing a plurality of mirrors that functions well in a broad spectral range from UV to NIR. Fig. 3 shows a schematic drawing of such a photo-induced DNA hybridization device. The DMD mirrors have two states when applying voltage onto them: "0" state at which the mirrors are tilted +10° and "1" state at which the mirrors are titled -10°. The incident light is directed to the light trap at the "0" state while the light goes to a lenslet array at the "1" state. The "0" and "1" states at each micromirror are determined by the mask pattern needed to correctly excite the nucleic acid hybridization process. The mask pattern chosen prevents the photo-excitation source from reaching particular nucleic acid molecules such that only desired immobilized nucleic acid molecules are subjected to the photo-excitation source. Masking is well-known in the art as illustrated by Pirrung, M.C., *Chem. Rev. 97:* 473-488 (1997) and U.S. Patent No. 5,744,101. The lenslet array directs the light beam on the semiconductor substrate, for example, to selectively induce the reactions. This unique feature allows the use of the optimum light source for photo-induced nucleic acid hybridization.

Transmissive-type SLM devices include a liquid crystal display and can be integrated with a nucleic acid memory device directly. Thus, the system may be more compact and simple in construction. However, most common electro-optical materials, such as liquid crystals, absorb UV light thus limiting their usage in the UV light wavelength region. Should the optimum wavelength for nucleic acid hybridization be in this wavelength range, then special developments are required to construct devices that are transparent in UV light wavelength range. PLZT, a solid-state material, is a significant potential candidate for an SLM device operating at UV wavelengths. *See* Fig. 4.

There are at least three methods of interfacing optics with aqueous systems - free space optical delivery system, guided wave optical system, and a combination of the two. Once the optimum photo-excitation conditions are found, the integration of a nucleic acid memory system with potential laser systems can be facilitated to provide the highest yield of photo-induced hybridization.

NSDOCID: <WO__0047600A1_I_>

Mobile Phase

The mobile phase may comprise an aqueous solution, e.g., a buffered salt solution or a solid or semi-solid phase such as a gel matrix, containing solubilized nucleic acid target molecules that are at least partial complements to the immobilized nucleic acid molecules. The mobile phase can be in direct contact with the photo-excitation source or immobilized array by solubilizing the target nucleic acid molecules into an agarose or a polyacrylamide gel or a gel containing PAA or PEO using methods well-known in the art. The mobile phase preferably comprises either at least partially complementary strands to all the immobilized nucleic acid sequences or long nucleic acid molecules to be stored by sequencing methods. However, the target strands have to be at least partially complementary for hybridization to occur. The target nucleic acid molecules in the mobile phase may have well-defined detectable labels, e.g., fluorescent, radioactive, phosphorescent, bioluminescent, chemiluminescent, biochemical or other labels that permit detection of the target nucleic acid molecules in the mobile phase or upon the substrate after hybridization, and should be stable over a broad range of thermodynamic and ambient light conditions. In addition, the target nucleic acid molecules should have sufficient mobility to diffuse to the immobilized nucleic acid molecules for hybridization to occur.

In order to create the most favorable conditions for the hybridization reaction with the immobilized nucleic acid molecules, the nucleic acid molecule concentration in the mobile phase should be maximized. Thus, the effects of pH, temperature, additives such as buffer and enzyme, and water content on the nucleic acid molecule concentration in the mobile phase should be determined as described below.

Nucleic Acid Selection

Nucleic acid molecules such as DNA, cDNA, RNA, mRNA and tRNA of at least eight nucleotides in length are useful in the invention. When UV light is employed, preferred nucleic acid molecules do not contain thymine or contain a small percentage of thymine. Additionally, nucleic acid molecules may contain a mismatch at every third base. Exemplary nucleic acid molecules is set forth in SEQ ID NOS:1 and 2.

Other nucleic acid molecules useful in the invention include nucleic acid molecules that have a melting temperature below the operating temperature of the photo-induced hybridization device. Methods of determining melting temperatures are well-known in the art.

Hybridization Detectors

Hybridization reactions between the immobilized nucleic acid molecules and the target nucleic acid molecules may be induced by light emitted at addressable locations. Information may be stored at specific addresses in the form of hybridized (double-stranded) polynucleotides. Retrieval of the data at each pixel may be made possible by means of distinguishing single- from double-stranded molecules. Possible methods for this purpose include tagging the target nucleic acid molecules with radioactive, fluorescent, phosphorescent, bioluminescent or chemiluminescent moieties, or another type of label which can be stimulated by an external light source and detected by, for example, a confocal detector or a charged coupled device (CCD) chip above the substrate. In either case, the object is to detect hybridized strands by a signal induced by illuminating the substrate with another light source. The confocal detector scans the substrate and collects data to form an image of the array. The CCD chip similarly provides an image of the array. When fluorescent labels are used, the fluorochrome on the single-stranded nucleic acid molecule preferably emits light at a different

3NSDOCID: <WO___0047600A1_I_>

wavelength and a lower fluorescence yield, than that on the hybridized double-stranded nucleic acid molecule.

Light can illuminate the target nucleic acid molecules and be collected effectively by the detector array due to the following: (I) the immobilized layer is a layer of immobilized single-stranded nucleic acid molecules. If the nucleic acid molecules are DNA, for example, then the lateral dimension of the nucleic acid molecules is about 9Å and the space between them is about 20Å. Therefore, most of the light can go through without being affected; (ii) the small dimension determines that no diffraction will occur; and (iii) the thickness of both the immobilized layer and the target layer are small and the total energy being absorbed is relatively small.

To facilitate the detection of hybridization on the substrate, the single-stranded nucleic acid molecules may be labeled with, for example, radioactive isotopes, fluorescent labels, phosphorescent labels, chemiluminescent labels, bioluminescent labels and enzyme labels. U.S. Patent No. 4,581,333 describes the use of enzyme labels to increase sensitivity in a detection assay. Radioisotopic labels are disclosed in U.S. Patent Nos. 4,358,535, and 4,446,237. Fluorescent labels (EP 144,914), chemical labels (U.S. Patent Nos. 4,582,789 and 4,563,417) and modified bases (EP 119,448) may also be employed. Fluorescent labels are most preferred. Specific fluorescent labels include digoxigenin via 3' end labeling with DIG-dUTP and fluorescein-12-dUTP or 6-FAM phosphoramidite. In the course of the analysis, the fluorochrome-labeled hybridized strands are excited by a light source such as laser and the emitted fluorescent light is detected to generate an image of the bands in the gel (Adleman, L.M., *Science 266*: 1021-1024 (1994)).

Labels may be employed prior to solubilization of the target nucleic acid molecules in the mobile phase. Alternatively, the immobilized nucleic acid molecules may be labeled. The nucleic acid molecules are labeled with at least one detectable label internally and/or at or near the 3' and/or 5' termini. Multiple labels, which may be the same or different, may be used.

As an alternative to labeling the nucleic acid molecules, the substrate, e.g., semiconductor, can serve as the light source for post-hybridization absorption studies, e.g., as a means of distinguishing hybridized nucleic acid molecules from unaltered nucleic acid molecules in the detection process.

Several methods have been suggested in the literature for detecting hybridization on chips (Smith, L.M., *et al.*, *J. Comput. Biol. 5:* 255-267 (1998)). These include electrical, linear and non-linear optical, and fluorescence methods. Specifically, options include Fourier transform infrared (FTIR) spectroscopy, surface plasmon resonance (SPR), and diffraction and spectroscopy methods.

Following the hybridization process, the nucleic acid fragments may be analyzed by gel electrophoresis wherein the fragments are separated according to their size (molecular weight). The original sequence can be read from the separated bands. Electrophoresis of nucleic acid molecules may be carried out in an agarose or polyacrylamide gel.

Detecting hybridization requires one to distinguish between hybridization on the substrate surface from hybridization in the mobile phase or from hybridization with neighboring pixels, and, second, to determine in which pixels hybridization has occurred. The success of this second step requires identifying the same pixels that were lased to within tolerable error limits. Errors may arise due to several factors. For example, there may be partial matching hybridization of target nucleic acid molecules in the mobile phase and diffusion of these molecules away from the pixel locations. There may also be hybridization at adjacent pixels due to poor laser registration or focusing, or from conduction/convection heating from the hybridization reactions in the designated pixel(s). Thus, the substrate may be prepared by exciting selected pixel locations with an external source such as, for example, laser light, to initiate the hybridization reactions in that area.

Verification of hybridization can be determined by performing experiments wherein the gel is stripped from the substrate so that the skilled artisan can work directly with the substrate surface nucleic acid molecules. Fluorescence and electrophoresis are reliable and standardized techniques that may be applied here.

3NSDOCID: <WO___0047600A1_I_>

The fluorescence experiments qualify the extent to which neighboring nucleic acid molecules are affected. Electrophoresis identifies the products of the hybridization, such as, for example, nucleic acid molecule experiment extensions by ligation reactions. Such experiments entail chemically or mechanically removing the nucleic acid molecules from the substrate.

Having shown that nucleic acid molecules on the substrate are hybridized under the action of laser light, the next task is to determine the hybridized nucleic acid molecule locations. As indicated above, several methods are known for detecting hybridization on substrates such as chips, including electrical, linear and non-linear optical, and fluorescence methods. Specifically, options include Fourier transform infrared (FTIR) spectroscopy, surface plasmon resonance (SPR), and diffraction and spectroscopy methods.

Hybridization Reactions

To induce hybridization, the light source, such as the VCSEL array, is focused onto each selected pixel of the immobilized nucleic acid array by a beam condensor/conditioner array such as a diffractive lenslet array. The photo-induced hybridization reactions take place on the immobilized array. The lenslet array may be fabricated from semiconductor materials such as, for example, glass, quartz, etc., depending upon the wavelength of the VCSEL array. For the UV range, preferred hybridization conditions include irradiation at about 300 nm with an intensity of about 14 mW/cm^2 from about 0.1 seconds to about 4.0 seconds. For the visible range, the preferred hybridization conditions include irradiation at about 430 to about 514 nm with an intensity of about 1.0 mW/cm^2 from about 0.4 seconds to about 30 seconds. For the NIR range, the preferred hybridization conditions include irradiation at about 920 nm with an intensity of about 4.45 mW/cm^2 from about 0.2 seconds to about 15 seconds.

The photo-excitation source initiates the following hybridization reactions:

nucleic acid (mobile phase) + nucleic acid (immobilized) = hybridized nucleic acid (1)
nucleic acid (mobile phase) + nucleic acid (mobile phase) = hybridized nucleic acid (2)

While the products of the reaction (1) will remain attached to the substrate surface, the products of reaction (2) will be present in the mobile phase. In addition, it is expected that, under certain conditions, photolysis will also lead to the formation of cyclobutane thymine dimers in the gel (discussed below). The nucleic acid molecules attached to the substrate are at the center of the proposed hybrid system, therefore reaction (1) is the focus of the invention. The product of reaction (2) is also hybridized nucleic acid molecules, but is present in the mobile phase instead of on the surface. These nucleic acid molecules may simply constitute a noise level in the system.

The complications that are raised due to reaction (2) and the dimerization reactions can be solved by vigorously washing the products of these reactions. Since these products are hosted in the mobile phase, aqueous-based solutions are excellent solvents that may be used for the washing process. Another technique to reduce the complications of reaction (2) and the dimerization reactions is the application of an external electrical field on the mobile phase since the nucleic acid molecules are negatively charged. This will force the hybridized nucleic acid in the mobile phase to migrate to the anode area.

Dimerization and Photo-Induced Hybridization

One of the objectives of the claimed invention is to achieve nucleic acid hybridization and crosslinking reactions of complementary strands while simultaneously minimizing dimerization reactions. UV-induced ligation and UV-induced covalent dimerization of oligonucleotides are well-known (Setlow, J. K., "The Molecular Basis of Biological Effects of Ultraviolet Radiation and Photoreactivation," In *Current Topics in Radiation Research,* Ebert M. & Howard,

INSDOCID: <WO___0047600A1_I_>

A. (eds.), North-Holland Publishing Co., Amsterdam 2:195-248 (1966); Rahn, R. O. and Patrick, M. H. "Photochemistry of DNA, secondary structure, photosensitization, base substitution, and exogenous molecules," in *Photochemistry and Photobiology of Nucleic Acids*, Wang S.Y. (ed.) Academic Press, New York, Biology 2:97-145 (1976); Carell, T., *Angew. Chem. Int. Ed. Engl. 34*:2491-2494 (1995); Love, J. D., *et al., J. Biol. Chem. 261*:10051-10057 (1986); Kypr, J., *et al., J. Biomol. Struct. Dyn. 11*:1225-1236 (1994)). UV light has been applied in DNA chemistry in a variety of applications despite the belief that its major effect is the creation of cyclobutane thymine dimers (Marmur, J., and Grossman, L., *Proc. Natl. Acad. Sci. U SA 47*:778-787 (1961); Rahn, R., *et al., Biophys. J. 9*:510-517 (1969); Glisin, V.R., and Doty, P., *Biochim. Biophys. Acta 142*:314-322 (1967); Patrick, M.H., *Photochem. Photobiol. 25*:357-372 (1977)).

Recently, Pospíšilová and Kypr (Pospíšilová, S. and Kypr, J., *Photochem. Photobiol. 65*:945-948 (1997)) have demonstrated UV light-induced crosslinking of DNA complementary strands. They found that the yield of crosslinked DNA was dependent on the DNA strand length and the UV light dose. Tiedge (Tiedge, H., *DNA Cell Biol. 10*:143-147 (1991)) has used UV light-induced cross-linking to immobilize target RNA molecules for *in situ* hybridization. Similarly, Affymetrix, Inc. (McGall, G.H., *et al., J. Am. Chem. Soc. 119*:5081 (1997); Anderson, R.C., *et al., Topics in Curr. Chem. 194*:117-129 (1998)) has used near-UV light for the synthesis of polynucleotide probe arrays on surfaces. In each of these applications, the results would have been hindered if dimerization were the controlling reaction, but the results were not hindered.

The issue of dimerization has also been addressed by Carell (Carrell, T., "Sunlight-Damaged DNA Repaired with Sunlight," *VCH Verlag GmbH*, Weinheim, *Angew Chem. Int. Ed. 34(22)*: 2491 (1995)). Carell has shown that the yield of photo-induced dimers depends on both the structure and sequence of the DNA as well as on the pH of the aqueous media. In the reversible reaction, a *cis-trans* cyclobutane thymine dimer forms upon UV light radiation (from about 200 to about 300 nm) of DNA.

Several steps can be taken to reduce the probability of formation of cyclobutane thymine dimers:

1 The backward reaction can be enhanced by using another radiation source in the range of from about 300 to about 500 nm.

2 The concentration of thymine can be reduced by selecting the sequences of nucleic acid molecules. Since the dimerization reaction is strongly dependent on the structure and sequence of the nucleic acid molecules, one may use single-stranded nucleic acid molecules with a little or no thymine. In this case, the hybridization reactions will depend more strongly on the guanine-cytosine pair (with three hydrogen bonds) than on the adenine-thymine pair (with two hydrogen bonds) Double helices with guanine and cytosine are more stable than those with adenine and thymine owing to the stronger hydrogen bonding.

3 Various quenchers can be used to suppress photo-induced dimerization It has been demonstrated that dimerization is virtually suppressed when the pyrimidine oligonucleotides $d(TC)_y$ or $d(C)_m$, are added to DNA carrying $d(TC)_x$ or $d(C)_n$ containing inserts, respectively (Lamichev, V.I., *et al.*, *Nature 344*:568-570 (1990))

4 Photo-induced dimerization is also pH-dependent owing to deprotonation reactions Thus, the pH factor will be used to further suppress the dimerization reactions

Characterization of Photo-Induced Hybridization Mechanisms and Kinetics

Characterization and optimization of photo-induced hybridization reactions require determination of the wavelength, intensity, pulse length, and pulse repetition frequency to maximize the yield of hybridized nucleic acid molecules. Such determinations may be made using mobiles phases such as aqueous solutions and gels and an external tunable laser or in semiconductor chips using an internal excitation source Analysis of the results may be conducted using classical

INSDOCID: <WO___0047600A1_I_>

fluorescence and electrophoresis experiments as well as advanced techniques for surface interrogation.

While not wishing to be bound by any particular theory, it is believed that the mechanism of the photo-induced hybridization reactions is based on photo-excitation that leads to the hydrogen bonding of the bases of two nucleic acid molecules as described above. The flash photolysis technique (described below) may be used to determine kinetic measurements, identify the chemical structure of the nucleic acid molecule in the excited state, and elucidate the reaction mechanism of the nucleic acid molecule and its moieties that can be applied in computational processes.

Flash photolysis is a powerful technique for studying transient species (mainly excited states and radicals) produced by photo-induced excitation of molecules using short pulses of UV and visible radiation. In this technique, the absorbance change induced in the sample solution by the UV or visible exciting flash is monitored by an analyzing light beam passing through the sample and reaching a detector (photomultiplier or diode) at the selected wavelength via a monochromator. The detector senses changes in the analyzing light intensity and converts them into electrical signals in order to display the temporal history of the absorbance in a computer. Laser flash photolysis experiments can determine the ability of light of different wavelengths (UV, visible, and NIR) to initiate hybridization reactions of nucleic acid molecules in aqueous solutions and gels and on solid surfaces, the most effective wavelength for initiating hybridization reactions, the effect of laser light on reaction rates, the mechanisms and kinetics of the hybridization process and the effects of temperature, concentration, pH, and additives, e.g., buffer and enzymes, on reaction rates. Results from such analyses can be used to control the reaction mechanism. For example, NIR flash photolysis may be used to measure the hybridization kinetics that cover the wavelengths from about 850 nm to about 970 nm.

Development of Chip Technology - Surface Chemistry of Nucleic Acid Molecules on Semiconductors

The experimental verification for demonstrating the self-assembly of nucleic acid molecules on a solid semiconductor substrate, such as silicon and GaAs, may be conducted as described by Ratner *et al.* (Boland, T. and Ratner, B.D., *Langmuir 10*:3845-3852 (1994); Park, D., *et al.*, *Macromolecules 28*: 2595-2601 (1995)). The morphological coplanar patterns of the self-assembled monolayer nucleic acid molecules may be achieved by either one or both of the following techniques: (1) using a deep UV photolithographic method to produce high-resolution spatial regions and (2) introducing an anchoring functional group onto the nucleic acid molecules to enhance the spontaneous self-assembly in a two-dimensional array on a solid substrate.

The use of a deep UV photolithographic method to produce high-resolution spatial regions depends on the photochemistry of the nucleic acid molecules and substrates. The mechanism of the surface photochemistry involves the identification of the nucleic acid molecule free radicals and their reaction kinetics. High-resolution spatial regions having different surface free energies for controlled adhesion of various nucleic acid molecules can be created using deep UV photolithographic techniques. It has been demonstrated that, upon deep UV irradiation, a properly chosen self-assembled monolayer (SAM) undergoes a photolytic cleavage reaction (Boland, T. and Ratner, B.D., *Langmuir 10*:3845-3852 (1994)). This principle is used to create spatially different areas of reactivity that may be amenable to re-modification with another SAM forming moiety. The ability to modify spatially the self-assembled monolayers may be important for high-density memory.

In the introduction of an anchoring functional group to nucleic acid molecules to enhance spontaneous self-assembly in a two-dimensional array on solid substrates, various anchoring groups may be introduced to the nucleic acid molecules to achieve better adhesion to the substrate. An anchoring functional group such as thiol or silane may be introduced to nucleic acid molecules to

RNSDOCID: <WO 0047600A1 I >

enhance the spontaneous self-assembly in a two-dimensional array on solid GaAs, for example. These functional groups strongly bind the nucleic acid molecules to the substrate surface thereby immobilizing it. The two-dimensional order is enhanced whenever the regular structure of the adsorbed molecule corresponds to the structure of the underlying substrate.

The mechanism of adsorption of DNA bases on to a gold surface has been studied by Boland and Ratner (Boland, T. and Ratner, B.D., *Langmuir 10*:3845-3852 (1994)) who have shown that the DNA bases self-assemble on crystalline gold in an ordered two-dimensional lattice. The process of self-assembling on gold was characterized by adsorption and surface migration which progressed in three stages. The initial appearance of random clusters of ordered molecules at the gold reconstriction sites was observed, i.e., some of the adsorbed molecules may stick to favorable, active sites on the gold while other diffusing molecules cluster with the ones initially adsorbed to build islands until the active site is completely occupied. The islands are clusters of molecules held in place by intermolecular force and adsorbed onto an active site on the gold. Only a limited number of molecules are allowed to be adsorbed onto active sites, occupation of all reconstruction sites by the clusters and rearrangement of the molecules to cover the remaining gold surface with a two-dimensional lattice. After step two is completed, the clusters have set up an epitaxial structure from which a wave of crystallization into a two-dimensional lattice can occur. The crystallization wave progresses perpendicular to the chains. When the substrate surface is GaAs, both the S-H group in thiol and the N-H group in the DNA bases may be functional groups. Scanning tunneling microscopy (STM) and electron spectroscopy for chemical analysis (ESCA) may be used to investigate the adsorption of nucleic acid molecules on GaAs.

Development of Chip Technology - Algorithm Development and Application

The use of nucleic acid in memory storage systems requires algorithm development and error analysis in order to reliably store and retrieve the data. Algorithm development and error analysis entails developing models of data storage using the unique properties of nucleic acid such as DNA. Exemplary memory models include simple ones derived by analogy with electronic memory devices to more complex models capable of massively parallel database searches. Models of other classical computing problems such as data sorting may also be employed Preferably, the memory storage devices are based on chips containing immobilized nucleic acid molecules in a gel of mobile target nucleic acid molecules. The chips are etched and the immobilized nucleic acid molecules are attached in addressable arrays. Addressable array chips are commonly used for DNA Sequencing by Hybridization (SBH) (Pease, A.C., *et al., Proc. Natl. Acad. Sci. USA 91*:5022-5026 (1994); Lipshutz, R., *et al., BioTechniques. 19*:442-447 (1995); Noble, D., *Anal. Chem. 67*:201A-204A (1995); Frutos, A.G., *et al., Nucl. Acids Res. 25:* 4748-4757 (1997)).

The concept for a nucleic acid memory chip is straightforward. There are two operations to be performed - storage and retrieval. There are specific difficulties associated with the storage and retrieval of addressable data in aqueous DNA solutions (Baum, E.B., *Science 268*: 583-585 (1995)), therefore, an addressable surface array is a desirable starting point for developing a more general three-dimensional storage system. Starting with surface-attached nucleic acid molecules, it may be possible take advantage of exotic topological structures that can be constructed with DNA (Chen, J. and Seeman, N.C., *Nature 350*:631-633 (1991); Seeman, N.C., *Accounts of Chem. Res. 30*:357-363 (1997); Winfree, E., *DIMACS Series in Discrete Mathematics and Theoretical Computer Science 27*:199-221 (1996)) in the form of three-dimensional memory storage systems.

The first model of nucleic acid memory to be considered is bit, byte, or word storage in direct analogy with digital electronic memory. When each pixel

RNSDOCID: <WO 0047600A1 I >

on the chip represents a single bit, then a chip with 4^{15} pixels, for example, could store 1 Gigabit or 128 Megabytes. This could be quite cumbersome, however, inasmuch as each pixel in a memory area would have to be queried to read the data. If the data is stored as bytes, rather than bits, then 2^8 (256) pixels are required to represent one byte. This reduces the number of pixels to be interrogated, but it reduces the chip capacity correspondingly, to a mere four Megabytes. Likewise, going to larger memory units, e.g., words, again reduces the available memory. However, word strategies minimize the noise level in the system and maximize the stability of hybridized complementary pairs.

In an alternative model, each pixel need only represent a 0 or 1 state. Therefore, fabrication of the chip is greatly simplified. In fact, all the pixels could contain the same nucleic acid molecules. Likewise, the composition of the target nucleic acid molecules is very simple. The disadvantages of this system are the tradeoffs in storage capacity versus the number of storage and retrieval operations needed, e.g., the number of pixels to be hybridized or fluoresced.

Another alternative model of nucleic acid memory is derived from DNA Sequencing by Hybridization (SBH). In this method, a nucleic acid fragment to be analyzed is incubated with a large set of oligonucleotide probes of known sequence. The extent to which the analyte nucleic acid molecule hybridizes with each probe is measured to identify those probes that perfectly complement the nucleic acid subsequences. This information can then be used to identify the analyte sequence. For example, this may be achieved by using detectably labeled, e.g., fluorescently-tagged, target nucleic acid molecules and reading the addresses of the labeled, e.g., fluorescing, pixels when the chip is illuminated. Recent experiments have illustrated hybridization to probes synthesized in 8-10 μm site with 4^{11} oligonucleotides placed in an array measuring 2-cm square. Array feature resolutions of about 1-2 μm are projected with advances in UV photolithography. Because of the combinatorial strategies used in fabricating the chip, the set of all 4^k oligonucleotides (k-mers) can be sequenced in 4 x k synthesis cycles (Pease, A.C., *et al., Proc. Natl. Acad. Sci. USA 91*:5022-5026 (1994)). The number of

unique probes increases exponentially with k while the number of steps increases only linearly.

The same concept can be applied to memory storage by encoding a data string in a base-4 code over the DNA alphabet set (A, C, G, and T). The data string is then a virtual DNA strand that can be stored on the chip by activating the pixels corresponding to its complementary subsequence. This assumes that the target medium has a complete set of fluorescence-tagged complementary subsequences to the probe strands. As an example, three bytes are stored in an array consisting of octanucleotide (8-mer) probes. *See* Fig. 5. Each byte requires a string of four characters from the base-4 alphabet set (A, C, G, and T). So three bytes translate to a 12-mer hypothetical DNA strand.

Ideally, only perfectly matched complementary polynucleotides hybridize. However, partial and slide matches are not only possible, but often cannot be avoided. Moreover, the stability of the hybridized nucleic acid molecules is dependent upon the percentage C/G content in the nucleic acid molecules so that even the perfect matches will vary in concentration (Smith, L.M., *et al.*, *J. Comput. Biol. 5:* 255-267 (1998)). More complex DNA word strategies have been suggested, (Frutos, A.G., *et al.*, *Nucl. Acids Res. 25:* 4748-4757 (1997)), but these lead to longer DNA strands and other complications associated with longer strands. Others have also pointed out that long strands are required to assure that all the nucleic acid molecules and their complements are distinct (Frutos, A.G., *et al.*, *Nucl. Acids Res. 25:* 4748-4757 (1997)). However, experimental results for SBH have shown that fluorescence signals from complementary nucleic acid molecules are 5-35 times stronger than those with single or double base-pair mismatches which demonstrates specificity in the identification of complementary sequences (Pease, A.C., *et al.*, *Proc. Natl. Acad. Sci. USA 91*:5022-5026 (1994)).

An alternative model of nucleic acid memory is one that exploits the massive parallelism that is well-known in other areas of DNA computing, particularly in the areas of combinatorial problems and data encryption. In this model, the chip is not used for writing to memory, but rather contains a permanent

database or library in which each pixel is encoded with different data. The database can be "searched" by populating the target medium with synthesized nucleic acid strands, preferably DNA strands, encoded with the search criteria. Only those sites containing the criteria will hybridize. The hybridization reactions may be initiated either by flashing the entire chip with laser light or by mixing an enzyme with the nucleic acid target strands. By illuminating the chip and using the fluorescence methods used in SBH, the results can be retrieved. For example, in searching for a fingerprint match, by encoding an unknown fingerprint into the base-4 alphabet set (A, C, G, and T) and synthesizing and replicating a DNA molecule to place in the target medium, the entire database can effectively be searched simultaneously (massive parallelism). Likewise, by using different fluorescence types, the search can be carried out for several pieces of data concurrently.

The following examples are illustrative, but not limiting, of the methods of the present invention. Other suitable modifications and adaptations of the variety of conditions normally encountered which are obvious to those skilled in the art are within the spirit and scope of the present invention.

Examples

Example 1 - Hybridization of an oligonucleotide in aqueous solutions induced by UV, Visible, and NIR light

Experimental Protocol

Materials and Solutions: An oligonucleotide with 27 bases (5'- CGA CGA CGA CGA CGA CGA CGA CGA CGA -3') (SEQ ID NO:1) was purchased in highly purified form (analytically estimated by HPLC and capillary electrophoresis) from Oligos Etc., Inc. (Wilsonville, OR). This particular sequence was used because an individual molecule has a low natural tendency for self-hybridization as every third base (A) is mismatched with a potential complementary strand (stem loop melting temperature of 0°C) and it does not contain thymine. *See* Fig. 14.

Thymine was excluded to prevent the thymine-thymine cyclobutane dimers upon irradiation with UV light.

Aqueous solutions containing 3.7 μM single stranded oligonucleotide (27 base pairs) were prepared at three pH values in different buffers: pH 4.5 (10 mM acetic acid), pH 7.8 (10 mM potassium phosphate), and pH 10.1 (10 mM sodium carbonate/sodium bicarbonate). All solutions were autoclaved at 121°C for 30 minutes prior to the addition of the oligonucleotides.

Irradiation: Photolysis was performed at several wavelengths. Irradiation at 450 nm and 920 nm was performed with an argon laser (Coherent Laser group, Santa Clara, CA) with an intensity of 4.45 mW/cm^2 at both wavelengths at exposure times of 0.2, 2.0, and 15 seconds. Irradiation at 300 nm was performed with an Hg lamp with an intensity of 14 mW/cm^2. Exposure times were 0.1, 0.6, and 4.0 seconds. The intensity of the visible source (430 nm) was 1.0 mW/cm^2, and exposure times were 0.4, 4.0, and 30 seconds. Laser irradiations were carried out at a pulse-width of 100 femtoseconds with an intensity of 4.0 mW/cm^2 at three wavelengths (450, 514, and 621 nm).

Analytical: The photolyzed samples were analyzed using precast polyacrylamide gels (15% polyacrylamide in TBE (tris-borate-EDTA)) from BioRad Co., (Hercules, CA). Samples (8.3 μL) were combined with 2 μL of gel loading solution (Sigma, type 1, gel loading buffer, 6X concentrate) prior to loading. The DNA ladder (Life Technologies, Inc.) (1.0 *ug/ul)* consisted of thirty-three 10 bp repeats and was suitable for sizing both single-stranded and double-stranded DNA fragments from 10 bp to 200 bp. The electrophoresis measurements were made with the ladder in the gel.

The gels were run in 1X TBE buffer (0.1 M Tris, 0.09 M boric acid and 1 mM EDTA, pH = 8) at low temperature (4°C) and low voltage (5 V/cm) to prevent denaturing of small DNA fragments by heat generated by the passage of electric current. DNA was visualized by ethidium bromide staining (UV transilluminator). Images were captured and stored via CCD camera (Eagle-Eye, Strategene). Band quantities were evaluated using NIH Image (v1.6). Error

analysis was performed on identical samples and on repeated experiments and the standard error was reported.

Results

Gel electrophoresis results for solutions irradiated at 300 nm (A), 450 nm (B), and 920 nm (C) are shown in Figs. 6A-6C. Gel electrophoresis images of the oligonucleotides before and after photo-induced hybridization in neutral aqueous solutions are shown. A strong band of the original 27 base oligonucleotide single-stranded is detected along with a weak slower-migrating band of a 54-base oligonucleotide. *See* Figs. 7A-7C. The quantity of dimer, as indicated by the intensity in the ethidium bromide stained polyacrylamide gels, is much higher in the photolyzed samples. For example, in samples photolyzed for 0.1 seconds at 300 nm, 27% of the original single-stranded oligonucleotide was converted to the higher molecular weight form. Similar results were obtained for other samples photolyzed for 15 seconds at 450 nm and 920 nm.

To examine whether the dimer was formed by hydrogen bond hybridization or by covalent bonding, the samples (9.0 µl) were heated to 95°C for 10 minutes to dissociate hydrogen bonds. After heating, only a trace of the dimer remained. Based on relative band intensities, the sample photolyzed for 0.1 seconds at 300 nm (Fig. 6A) showed that 26% of the original single-stranded oligonucleotides had been converted into the higher molecular weight form, but after heating, only a trace of dimer remained (approximately 3%). *See also* Fig. 7A. Similar results were obtained for other samples photolyzed at 450 and 920 nm (Figs. 6B and 6C). *See also* Figs. 7B and 7C. Additionally, 8 μl samples were treated with 7 μl NaOH solution (pH 13) which resulted in complete dissolution of the higher molecular weight band.

To examine whether the intensities of the bands in Figs. 7A-7C increase linearly with the weight of the oligonucleotide, various concentrations of the oligonucleotide were run in the same gel. For each concentration a control

ISDOCID: <WO___0047600A1_I_>

(unexposed) sample was also processed on the gel. The sum of the intensities of the two bands, the 27- and 54-mer bands, was plotted against the weight of the oligonucleotide. A typical plot is shown in Fig. 8. In all cases, the band intensity increased linearly with oligonucleotide weight up to an intensity of 8×10^3. Therefore, any band intensity higher than this value was discarded.

In addition, DNA melting experiments were performed. The absorbance at 260 nm of 100 μl of photolyzed samples (pooled from several experiments) was measured between 20°C and 90°C in 5-20°C increments (Figs. 9A and 9B). Aliquots were removed at each temperature and analyzed by polyacrylamide gel electrophoresis. As the temperature increased, the upper band intensity decreased and the solution A_{260} increased monotonically, which is typical for DNA hybrid melting (A_{260} ss$\approx$1.5 A_{260} ds). The theoretical melting temperature for a perfect match 27mer based on similar %GC content is 83.8°C (Oligos Etc. Inc., Wilsonville, OR). As noted in the Experimental protocol section, the oligonucleotide used here was constructed for third base pair mismatch when two single-stranded oligonucleotides hybridize, hence the theoretical melting temperature for the 27mer employed is 40-50°C. The observed melting temperature ($\approx$60°C) is bounded by these extremes. These results constitute firm evidence that photo-induced hydrogen bond dimerization was achieved rather than covalent bonding.

Table 1 shows a calculated "efficiency" of the UV (300 nm) photo-induced hybridization at pH $\approx$7.8 and at various oligonucleotide concentrations. This efficiency is defined as the intensity of the upper band (dimer) divided by the sum of the upper and lower band intensities. This result was only recorded for samples shown to be in the linear range of the calibration solutions and the gels (quantity < 160 ng per band). Namely, the intensity of the ethidium bromide signal of a single band was linear with 4.0g DNA up to 8×10^3 intensity units. Data from lanes with intensities above this limit were not included. At 300 nm, an overall efficiency of 27 $\pm$ 3.2% was found after photo-induced hybridization, compared with 10.6 $\pm$ 2.5% in the unirradiated controls.

Table 1

Efficiency of Photo-Induced Hybridization of Oligonucleotides at Various DNA Concentrations

Sample pH = 7.8	Amount of DNA (ng)	Upper Signal	Lower Signal	Total Signal	Efficiency (%)	
Photolyzed (0.1 sec @ 300 nm)						
1	250	2306	5987	8293	27.8	
2	250	2542	7143	9685	26.3	
3	129.6	1441	5123	6564	22.0	
4	97.5	1163	2953	4116	28.3	Avg.: 27.0±3.2
5	54.9	841	1913	2754	30.5	
Unirradiated sample[a]						
1	250	898	6742	7640	11.8	
2	250	622	7155	7777	8.0	
3	250	1164	7810	8974	13.0	
4	250	965	7566	8531	11.3	
5	250	696	6431	7127	9.8	
6	129.6	687	6353	7040	9.8	
7	129.6	451	4712	5163	8.7	Avg.: 10.6±2.5
8	129.6	662	5904	6566	10.1	
9	91.5	560	4271	4831	11.6	
10	91.5	385	4176	4561	8.4	
11	91.5	385	5045	5430	7.1	
12	54.9	411	2162	2573	16.0	
13	54.9	312	3091	3403	9.2	
14	54.9	576	3553	4129	14.0	

[a]data are obtained from three different gels.

Additional experiments were run at various pH and wavelengths and hybridization ratios were calculated. Specifically, each sample (control and irradiated) contained the same amount of DNA and the bulk of the DNA was in the single-stranded form. The intensity of the upper band (dimer) from irradiated samples was divided by the intensity of the upper band from controls (after

normalization by lower band intensity) because this ratio between double-stranded hybrids is a more sensitive indicator of the effect of photolysis. In Fig. 10, the effect of pH on the HR (300 nm, exposure time of 0.1 seconds) is shown. This was not surprising as pH extremes leave ionic interactions more unlikely due to the predominance of protonated or deprotonated forms. Also, in Fig. 10, the relative HRs for all samples at three wavelengths are shown. Although UV (300 nm) irradiation was most effective, significant hybridization was formed in the NIR range (920 nm). The values of the hybridization ratios were found to be 3.5 and 2.5 at 300 nm and 920 nm, respectively.

The effect of pH on the hybridization ratio at 300 nm with an exposure time of 0.1 seconds was investigated. The hybridization ratio decreased from about 3.0 at a pH of about 7.8 to about 1.0 at a pH of about 4.5 and about 0.6 at a pH of about 10.1. Fig. 11 shows the effect of the exposure time with 300 nm light on the hybridization yield of 37 μM at pH 7.8. Hybridization yield increased to a maximum value at 0.1 seconds. At higher exposure times (0.6 and 4.0 seconds), UV photolysis appears to induce scission of the hybridized molecules. Moreover, the relative hybridization ratios of the DNA sample (pH= 7.8) photolyzed with a broad spectrum band which includes UV, VIS and IR regions are shown in Fig. 12. In these experiments, both, steady state and flash photolysis techniques were utilized. As expected, UV irradiation (300 nm) induced more hybridization than the visible and IR irradiation.

Overall, photo-induced hybridization was successfully achieved. The results show that irradiation of a single-stranded DNA by light with wavelength of 300, 450, and 920 nm, induces hybridization between complementary bases. Covalent dimerization was prevented by the absence of a thymine moiety. The photolysis improved the hybridization reaction especially in the vicinity of pH of 7.8. That is, aqueous solutions of DNA with pH of about 7.8, irradiated by UV at 300 nm for 0.1 seconds, gave rise to the highest relative extent of hybridization.

Discussion

The results demonstrate considerable hybridization of the oligonucleotides as a result of UV irradiation. Moreover, visible and NIR irradiation led to hybridization, although to a lesser extent. Since hybridization involves the dissociation of hydrogen bonds (between a base and water molecule) and the formation of other hydrogen bonds (between complementary bases), all the wavelengths used in this study had sufficient energy to effect such changes.

Since UV light is absorbed by the DNA bases, an additional mechanism appears to be operating in this case as a result of direct excitation of the oligonucleotide. The excitation is absorbed by the oligonucleotide and the excited state can undergo charge separation to produce an anion radical and a cation radical of different bases. It is known that radiolysis and photolysis of DNA produce cation radicals, mainly $A^{\bullet+}$ and $G^{\bullet+}$, and anion radicals, mainly $C^{\bullet-}$ and $T^{\bullet-}$, due to the relative electron affinities of the DNA bases. Steenken (Steenken, S., *Free Rad. Res. Commun. 16*:349-379 (1992); Steenken, S., *Biol Chem. 378*:1293-1297 (1997)) showed that the anion radicals react rapidly with protons and the cation radicals can readily donate protons. As a result, if $C^{\bullet-}$ of one oligonucleotide approaches $G^{\bullet+}$ of another nucleotide, rapid proton transfer between them will take place to produce two neutral radicals that are hydrogen-bonded. These radicals are eventually repaired, for example, by charge migration and recombination, but the repaired bases can remain hydrogen bonded. The findings that UV-induced hybridization is more effective at pH 7.8 than at pH 4.5 or pH 10.1 may be interpreted to support this mechanism, i.e., when the medium is more likely to supply the proton to $C^{\bullet-}$ (at low pH) or more likely to remove a proton from $G^{\bullet+}$ (at high pH), then the probability of proton transfer between these two species is diminished and thus the extent of hybridization is decreased.

With visible and NIR light, for which charge separation is unlikely, the principal path whereby the hybridization occurs remains unclear. As suggested above, a reasonable postulate is photon-induced breakage of hydrogen bonds

INSDOCID: <WO___0047600A1_I_>

between the nucleotide and water molecule followed by the exothermal hydrogen bonding of complementary bases. Despite the higher efficiency of UV light in effecting hybridization, it is more practical to use visible or NIR light sources for analytical and/or computer applications because of the avoidance of covalent (irreversible) linkage. Also, the observation that substantial hybridization occurred upon 920 nm irradiation is noteworthy as GaAs emission is typically at 920 nm. Thus, a possible memory device could contain a gel of oligonucleotide strands attached to a semiconductor wafer by appropriate end groups as described in detail above. Light emission at selected sites would cause hybridization at these sites. The extent of hybridization may be increased above the levels observed in this example by selecting oligonucleotide sequence and length (match/mismatch, GC content) (Ikuta, S., *et al., Nucleic Acids Res. 15*:797-811 (1987)). Also, the oligonucleotide concentration could be much higher for gene array and computer applications and the increased proximity of complementary strands may enhance hybridization.

The energy of a DNA hydrogen bond is typically about 0.35 eV. There is sufficient energy to dissociate one such bond not only by an ultraviolet (UV) photon (typically about 4.0 eV), but even by a near-infrared (NIR) photon (from about 1.0 to about 1.3 eV) Thus, even GaAs emission (about 1.3 eV) can, in principle, dissociate an adjacent pair of hydrogen bonds, thereby raising immediate prospects of erasing stored information.

Example 2 - NIR-Induced Hybridization

In these experiments, aqueous solutions of oligonucleotides (5' TGCA TGCA TGCA TGCA TGCA TGCA TGCA Y) (SEQ ID NO:2) were irradiated by NIR at 850 nm with gallium aluminum arsenide (GaAlAs) (Spectra Diode Laboratories) with the intensity of 76.4 W/cm^2. The sequence of this oligomer is self-complementary with no mismatches (dimer melting temperature 77.9°C). It also forms a stem loop structure (melting temperature 120°C based on the GC content).

Preliminary results demonstrate that irradiation with an 850 nm laser results in two distinct bands on polyacrylamide gels. The intensity of these bands varies with the type of buffer used for diluting the oligonucleotide. The following standard techniques were used to demonstrate that the upper band is indeed a double-stranded oligomer composed of two single-stranded hydrogen bonded oligonucleotides: (1) boiling the sample (5 minutes at 95°C) caused the upper band to disappear; (2) increasing the pH with NaOH (pH 13) also caused the upper band to disappear; (3) adding S I nuclease (which attacks single-stranded DNA specifically) significantly reduced the brightness of the lower band; and (4) measuring the UV absorption (260 nm) of the oligomer showed a 23% increase in absorption as the temperature was raised from 40°C to 90°C. The DNA melting curve determines the transition from double-stranded DNA into single-strand as shown in Fig. 13.

This invention has been described in specific detail with regard to specific materials and methods for the photo-induction of nucleic acid hybridization and applications thereof. Except where necessary for operability, no limitation to these specific materials is intended nor should such a limitation be imposed on the claims appended hereto. From the foregoing description, one skilled in the art can easily ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions without undue experimentation. All patents, patent applications and publications cited herein are incorporated by reference in their entirety.

What Is Claimed Is:

1. A nucleic acid memory device comprising a photo-excitation source, a substrate with nucleic acid molecules immobilized thereto and attached and in a combinatorial array, a mobile phase containing nucleic acid molecules at least partially complementary to the immobilized nucleic acid molecules, and a detector which detects hybridization within the combinatorial array.

2 The nucleic acid memory device of claim 1, wherein the photo-excitation source and the substrate are the same.

3. The nucleic acid memory device of claim 1, wherein the photo-excitation source is a semiconductor.

4. The nucleic acid memory device of claim 1, wherein the substrate is a semiconductor

5. The nucleic acid memory device of claim 1, wherein the immobilized nucleic acid molecules are DNA molecules and the target nucleic acid molecules are DNA molecules.

6. The nucleic acid memory device of claim 1, wherein the detector is selected from the group consisting of: a confocal detector, a charged coupled device, and a flourescent tag.

7. The nucleic acid memory device of claim 6, wherein the flourescent tag is selected from the group consisting of: DIG-dUTP, fluorescein-12-dUTP, and 6-FAM phosphoramidite.

BNSDOCID: <WO 0047600A1 I >

8. The nucleic acid memory device of claim 1, further comprising a computer.

9. A method of inducing nucleic acid hybridization comprising:

(a) attaching a monolayer of single-stranded nucleic acid molecules in a combinatorial array of pixels to a substrate, wherein the nucleic acid molecules are immobilized;

(b) contacting the substrate with a mobile phase containing one or more target single-stranded nucleic acid molecules at least partially complementary to the immobilized nucleic acid molecules; and

(c) applying a photo-excitation source onto each pixel of the combinatorial array to induce hybridization between the immobilized nucleic acid molecules and the one or more target nucleic acid molecules.

10. The method of claim 9, wherein the photo-excitation source is a semiconductor.

11. The method of claim 10, wherein the photo-excitation source is a VCSEL array.

12. A method of inducing hybridization of two or more nucleic acid molecules which are at least partially complementary, comprising applying a photo-excitation source to the two or more nucleic acid molecules, whereby hybridization is induced.

13. A method for amplifying at least one double-stranded nucleic acid molecule, comprising:

(a) providing a first and second primer, wherein the first primer is complementary to a sequence at or near the 3'-terminus of the first strand of the nucleic acid molecule to be amplified and the second primer is complementary to a

sequence at or near the 3'-terminus of the second strand of the nucleic acid molecule to be amplified;

(b) mixing the first and second primers, and the nucleic acid molecule to be amplified, with one or more polypeptides having nucleic acid polymerase activity thereby forming a mixture;

(c) illuminating the mixture with light under conditions favoring the photo-induced hybridization of the first primer to the first strand of the nucleic acid molecule, and the second primer to the second strand of the nucleic acid molecule;

(d) producing a third nucleic acid molecule complementary to the first strand and a fourth nucleic acid molecule complementary to the second strand;

(e) denaturing the first and third strands, and the second and fourth strands; and

(f) repeating steps (a) to (e) one or more times.

14. A method of sequencing a nucleic acid molecule, comprising:

(a) mixing a primer with a first nucleic acid molecule to form a first mixture;

(b) illuminating the first mixture with a photo-excitation source under conditions favoring the photo-induced hybridization of the primer with the first nucleic acid molecule;

(c) mixing the first mixture with one or more deoxyribonucleoside triphosphates (dNTPs), one or more polypeptides having nucleic acid polymerase activity and one or more terminator nucleotides thereby forming a second mixture;

(d) incubating the second mixture under conditions sufficient to synthesize a random population of DNA molecules complementary to the first nucleic acid molecule, wherein the synthesized nucleic acid molecules are shorter in length than the first nucleic acid molecule and wherein the synthesized nucleic acid molecules comprise a terminator nucleotide at their 3' termini; and

BNSDOCID: <WO_ _0047600A1_I_>

(e) separating the synthesized nucleic acid molecules by size so that at least a part of the nucleotide sequence of said first nucleic acid molecule can be determined.

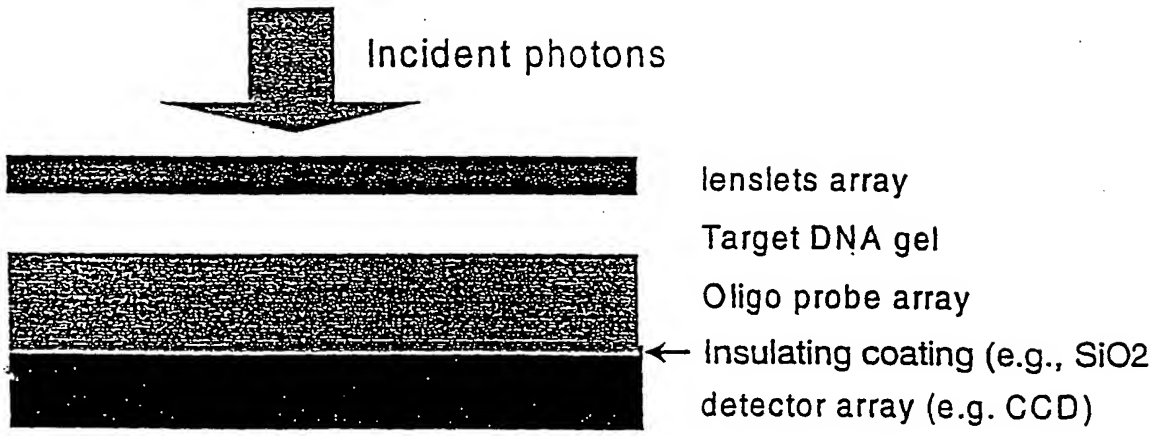
15. The method of claim 14, wherein the one or more terminator nucleotides is selected from the group consisting of: ddTTP, ddATP, ddGTP, ddITP, ddCTP and mixtures thereof.

16. A method of inducing nucleic acid hybridization comprising:

(a) obtaining a substrate comprising a monolayer of single-stranded nucleic acid molecules in a combinatorial array of pixels immobilized thereon and a mobile phase containing one or more target single-stranded nucleic acid molecules at least partially complementary to the immobilized nucleic acid molecules; and

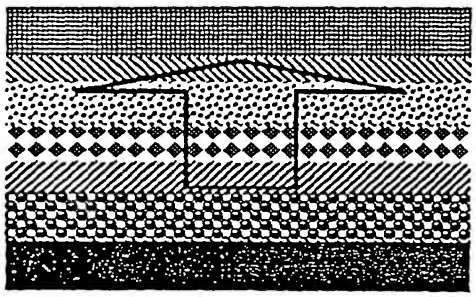
(b) applying a photo-excitation source to at least one pixel of the combinatorial array to induce hybridization between the immobilized nucleic acid molecules and the one or more target single-stranded nucleic acid molecules.

Figure 1



INSDOCID: <WO___0047600A1_I_>

Figure 2



Detector array (e.g. CCD array)

Beam condensor (Lenslet array)
Target DNA gel

Oligo probe array
Beam conditioner (e.g., Lenslet array)
VCSEL array (e.g., GaN VCSEL array)

Substrate

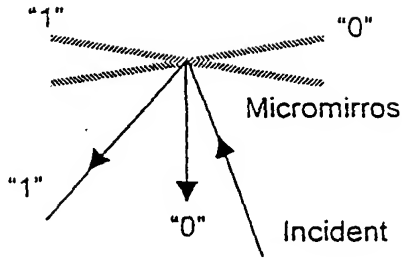


Figure 3A

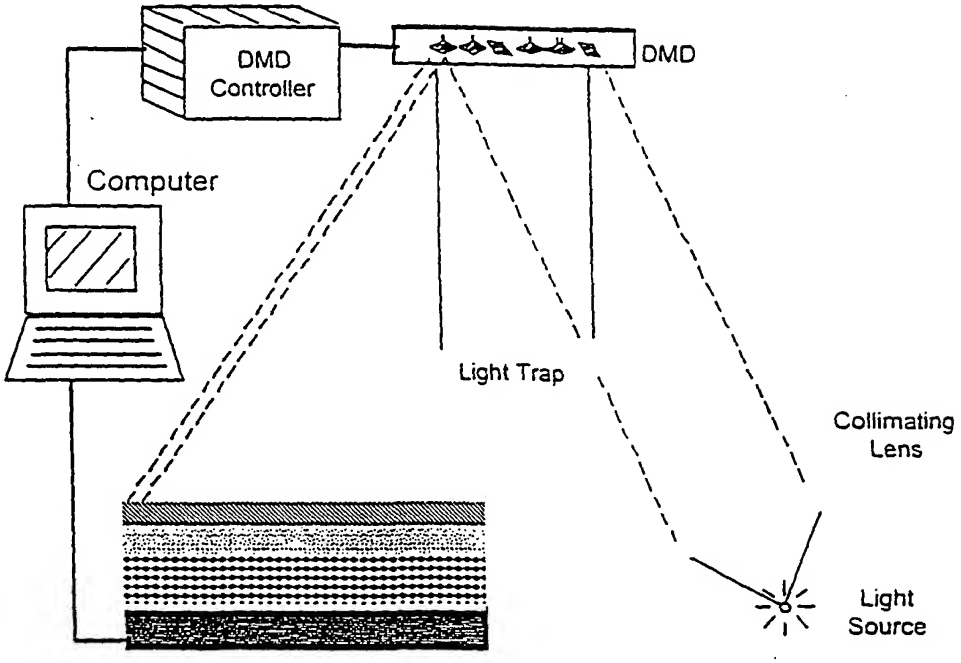


Figure 3B

Fig. 4

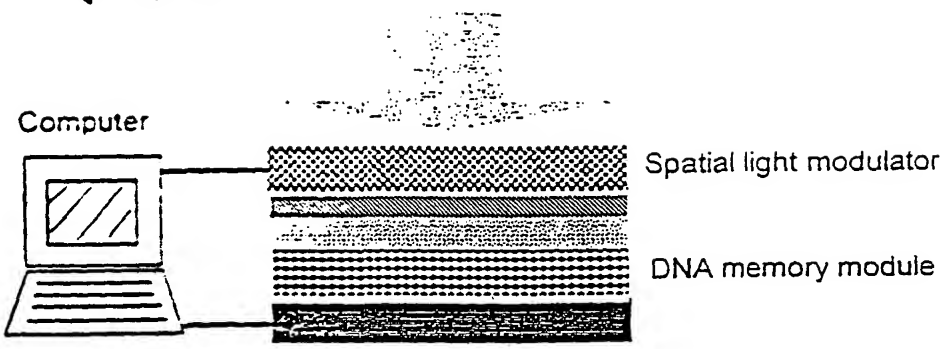
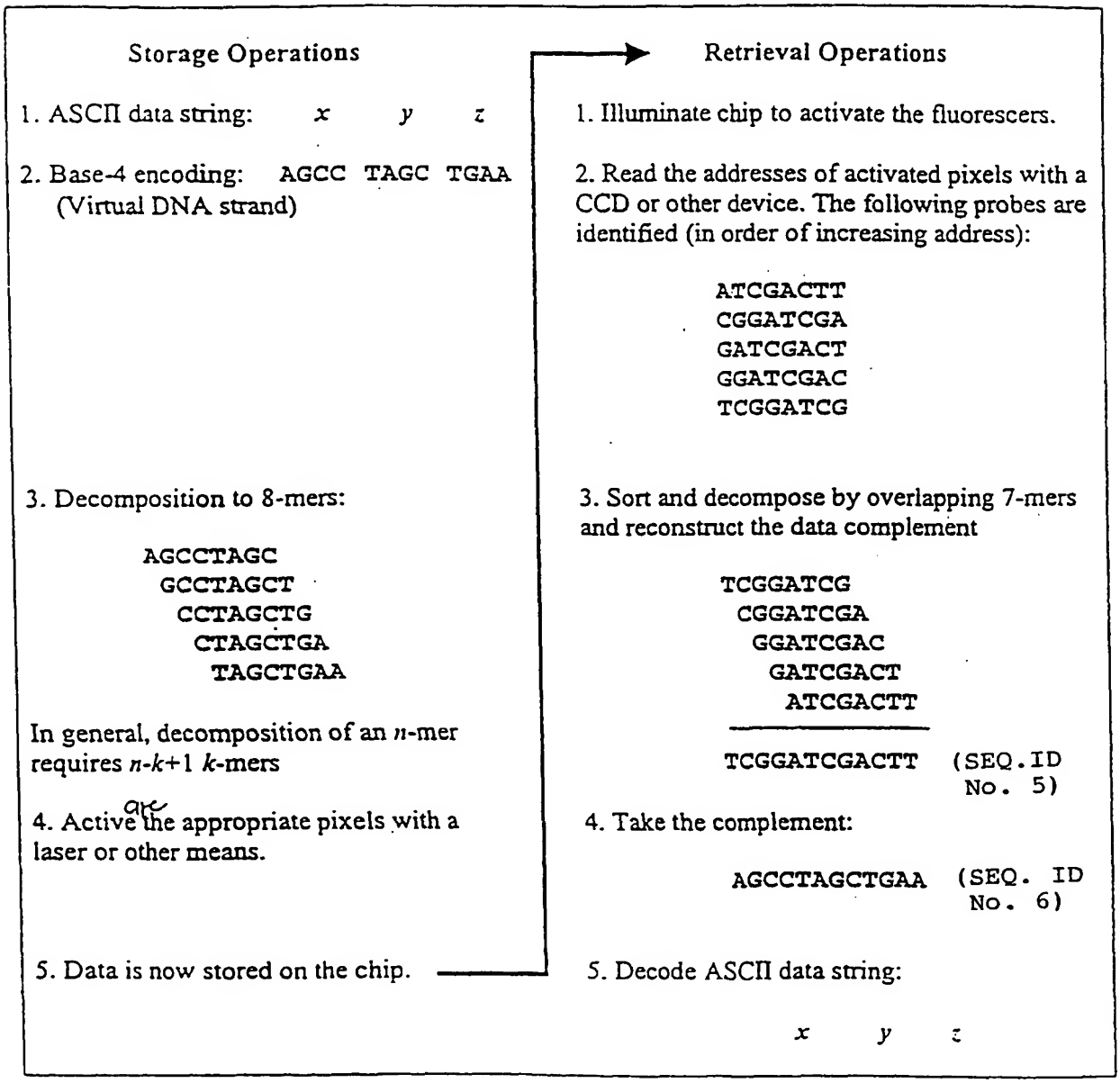


Figure 5

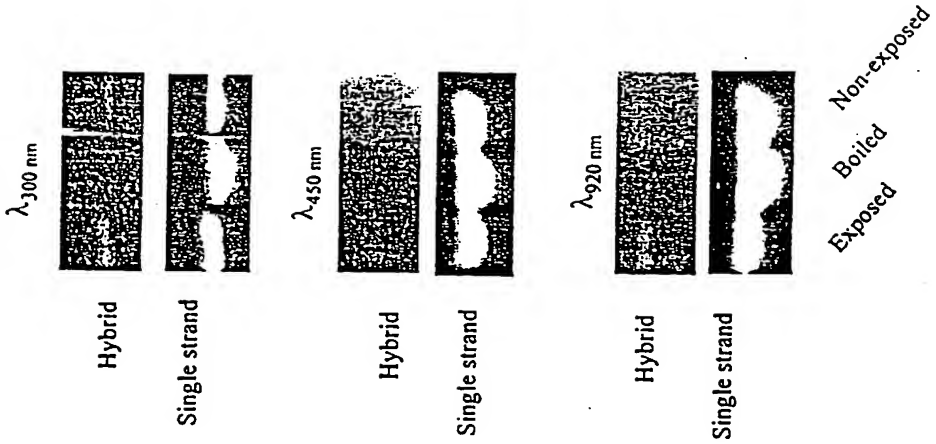


3NSDOCID: <WO___0047600A1_I_>

Figure 6A

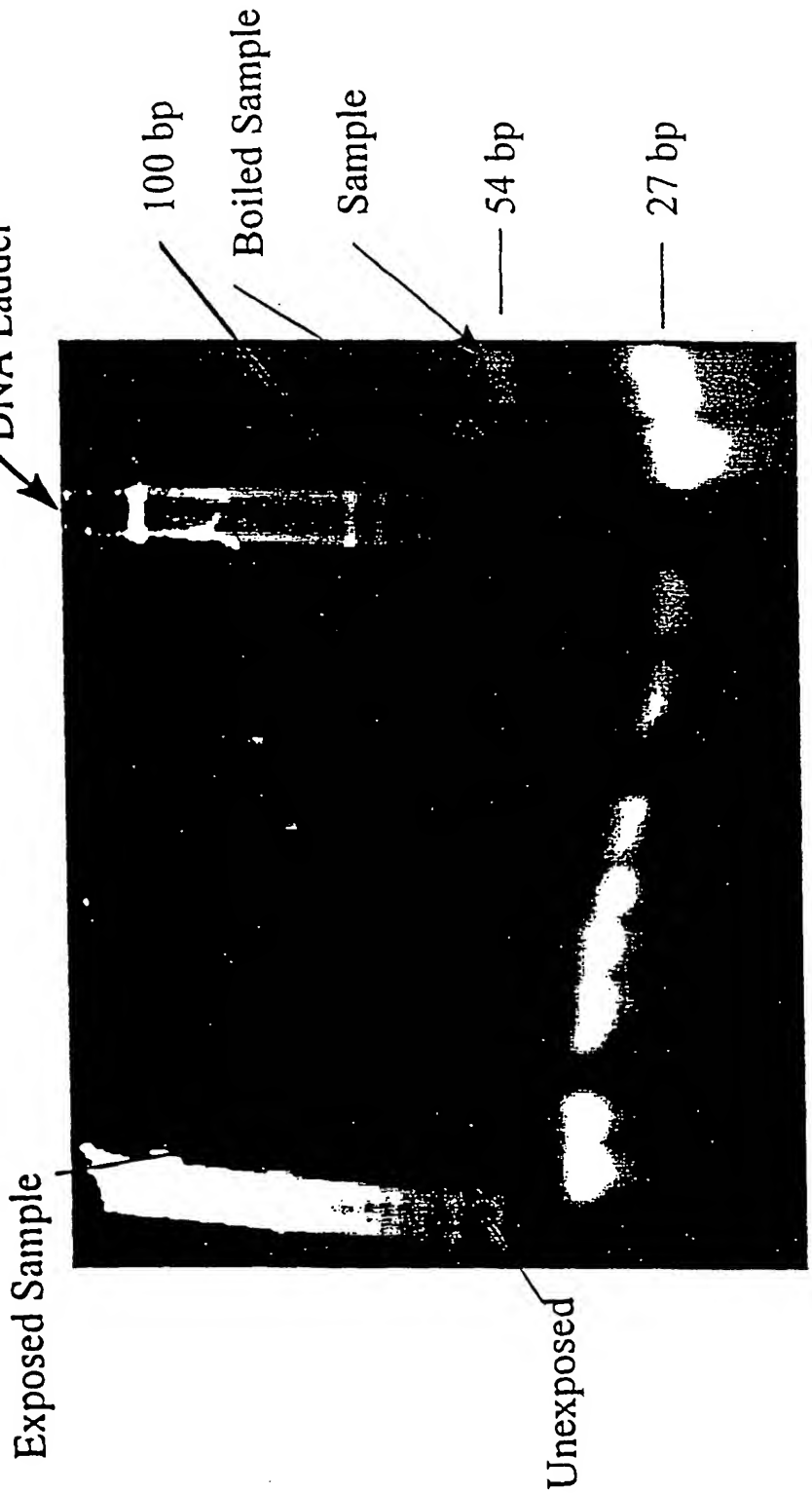
Figure 6B

Figure 6C



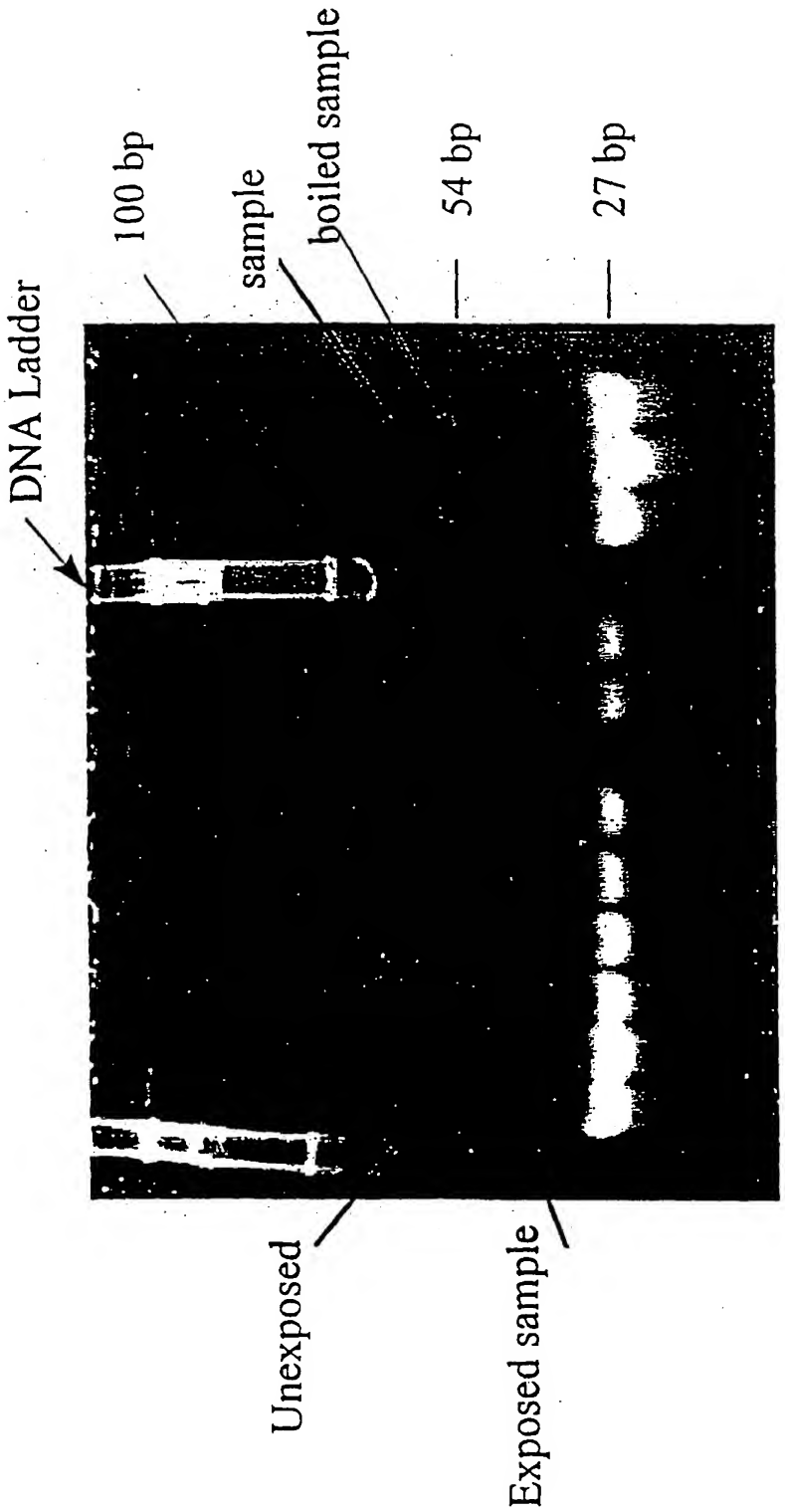
NSDOCID: <WO___0047600A1_I_>

Figure 7A



3NSDOCID: <WO___0047600A1_I_>

Figure 7B


DNA Ladder
100 bp
sample
boiled sample
54 bp
27 bp
Unexposed
Exposed sample

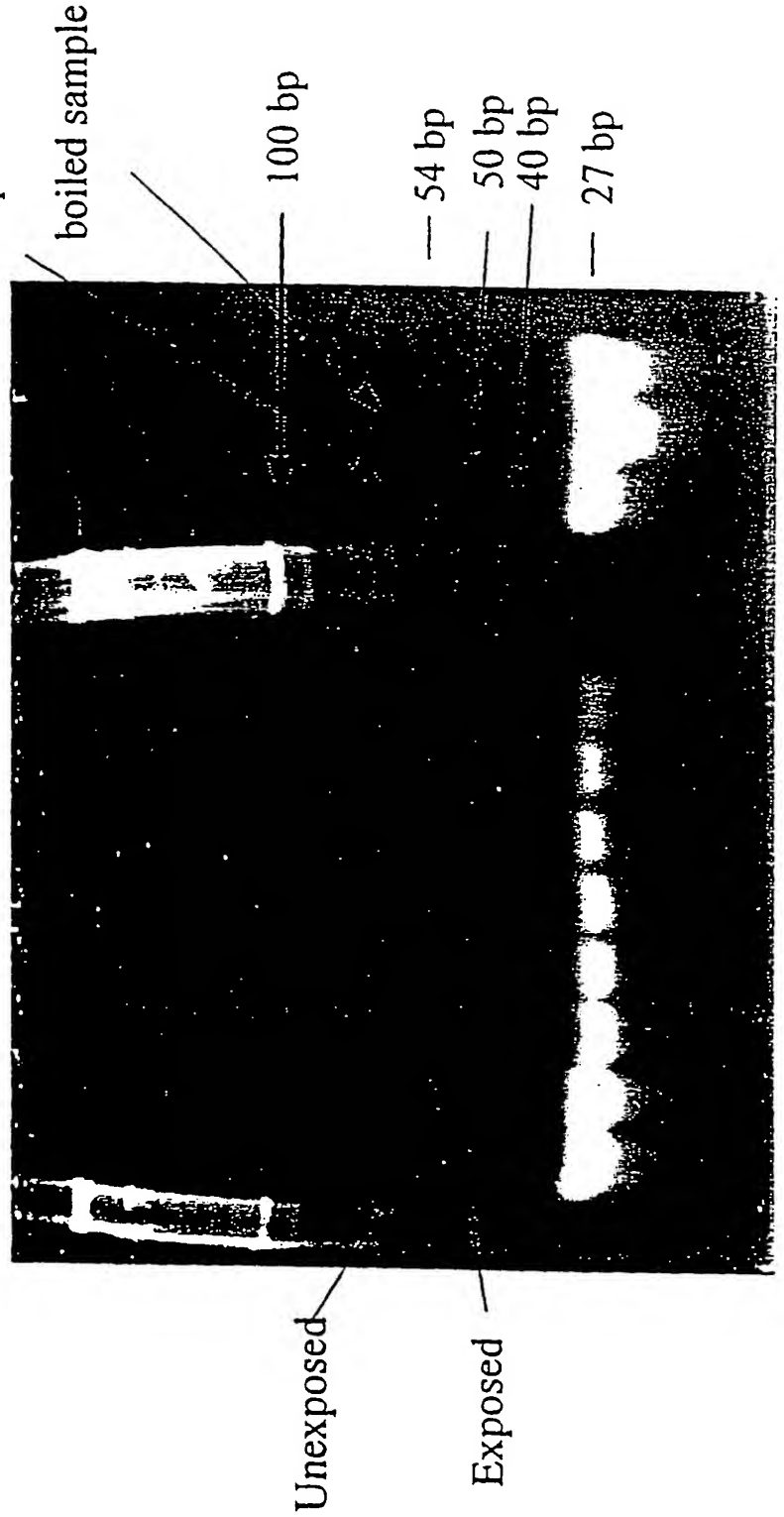
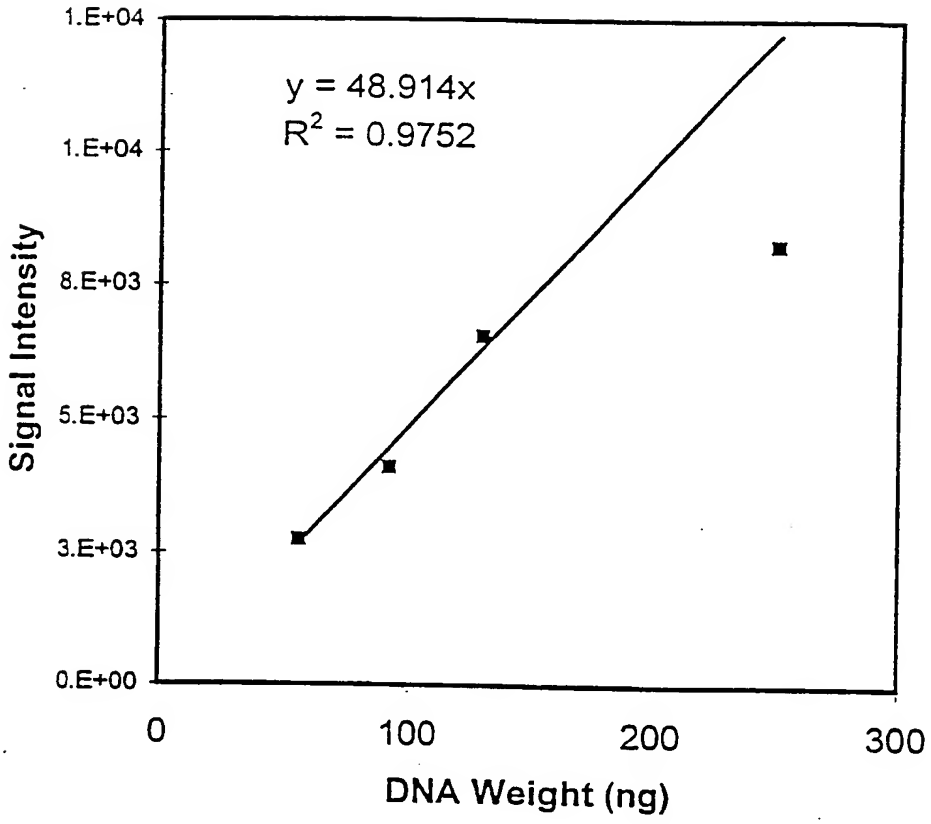

sample
boiled sample
100 bp
54 bp
50 bp
40 bp
27 bp
Unexposed
Exposed

Figure 7C

Figure 8



INSDOCID: <WO___0047600A1_I_>

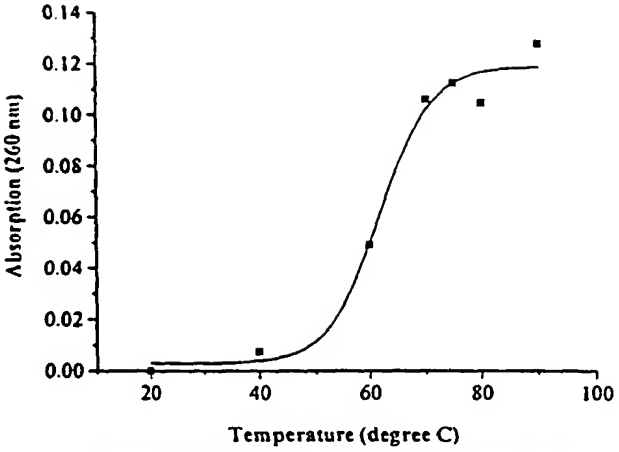


Figure 9A

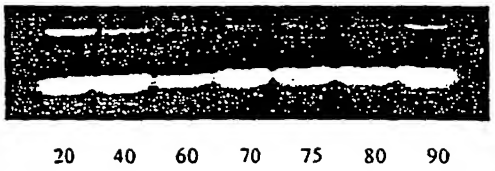


Figure 9B

INSDOCID: <WO___0047600A1_I_>

Figure 10

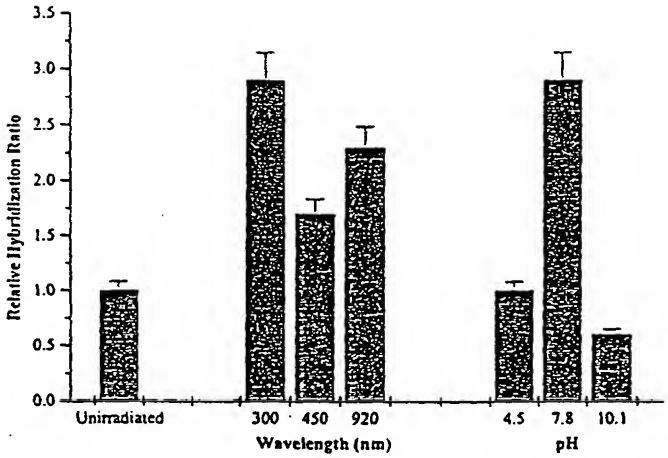
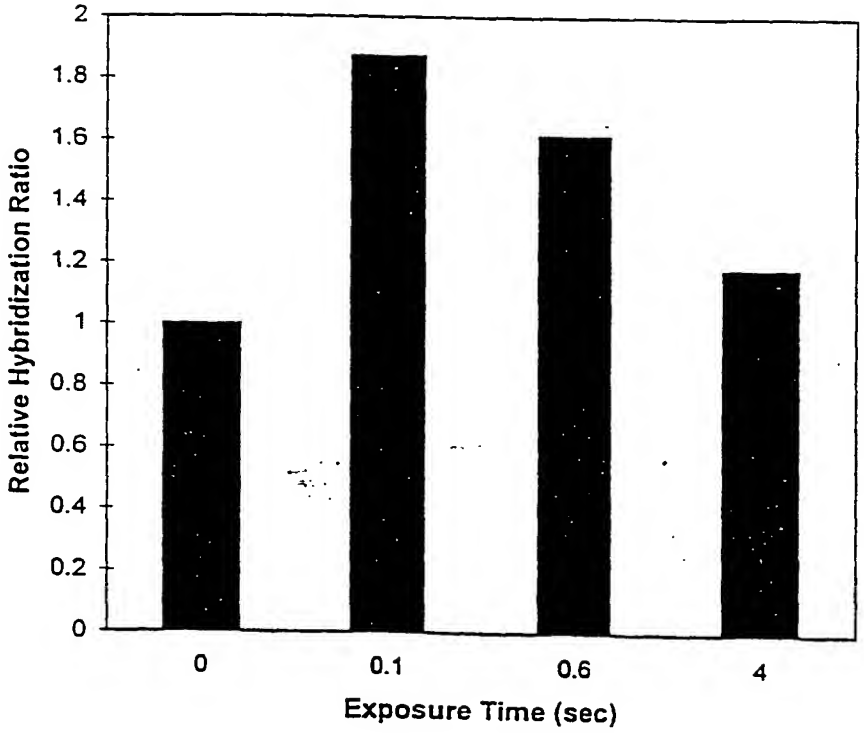
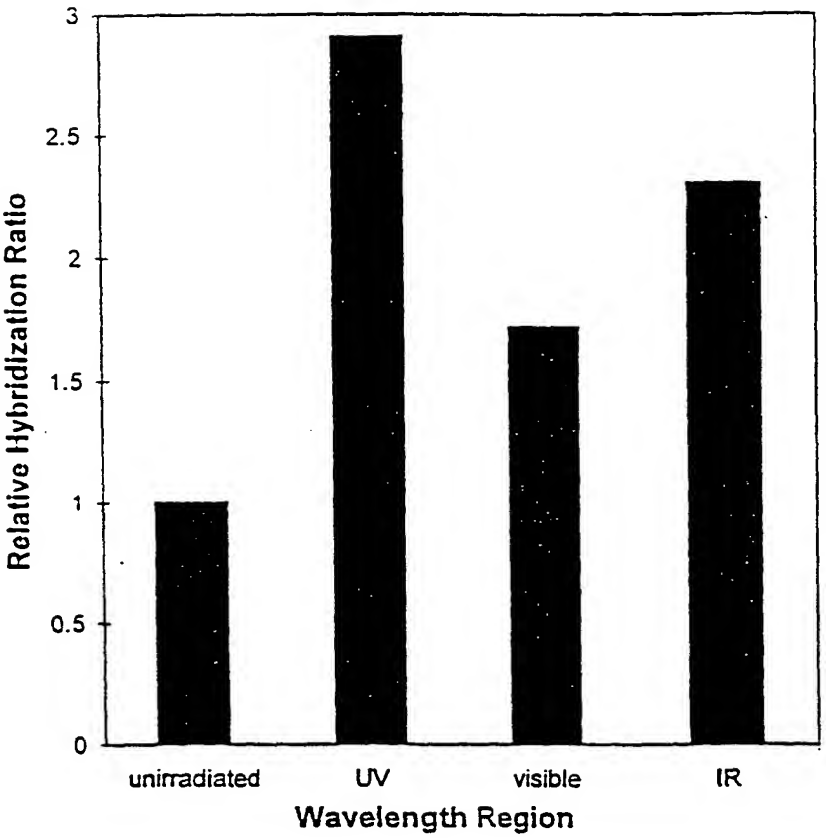

Relative Hybridization Ratio
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Unirradiated
300
450
920
Wavelength (nm)
4.5
7.8
10.1
pH

Figure 11



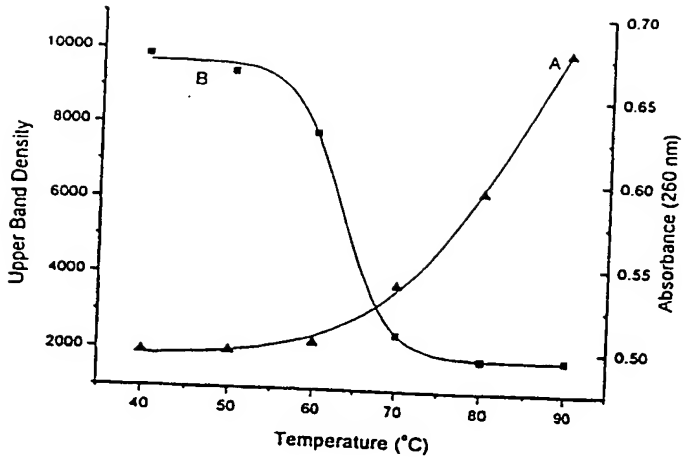
ENSDOCID: <WO___0047600A1_I_>

Figure 12



WSDOCID: <WO___0047600A1_I_>

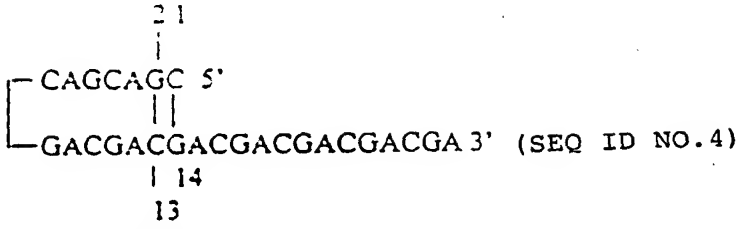
Figure 13



NSDOCID: <WO___0047600A1_I_>

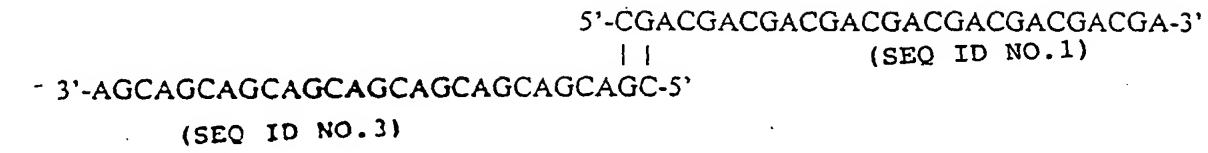
Figure 14

Stem Loop Structure:



G = 0.7 kcal/mol
Loop Tm = 0 C

Unstable Homodimer:



NSDOCID: <WO___0047600A1_I_>

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/03357

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) :C07H 21/00, 17/00; C12Q 1/68
US CL :536/24.3; 435/6
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/24.3; 435/6, 5, 91.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST, STN, MEDLINE, BIOSIS, CAPLUS, EMBASE, GENBANK
search terms: hybridization, sequencing, irradiation, photo-induction, nucleic

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98/12559 A1 (DEMERS) 26 March 1998, see entire document.	1-8
X	US 5,565,322 A (HELLER) 15 October 1996, see entire document.	1-8
---		----
Y		9-16
X	US 5,412,087 A (MCGALL et al) 02 May 1995, see entire document.	1-8
---		----
Y		9-16
Y	US 5,002,867 A (MACEVICZ) 26 March 1991, see entire document.	1-16
Y	US 5,795,722 A (LACROIX et al) 18 August, 1998, see entire document,	14-16

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
27 MARCH 2000	07 JUN 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer ARUN CHAKRABARTI Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)*

NSDOCID: <WO___0047600A1_I_>

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/03357

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PEASE.A.C. et al. Light-generated Oligonucleotide Arrays for Rapid DNA Sequence Analysis. Proc. Natl. Acad.Sci. USA. 04 January 1994, Vol. 91. pages 5022-5026, see entire document.	1-16
X ---- Y	US 4,599,303 A (YABUSAKI et al) 08 July 1986, see entire document, especially column 5, line 8 to column 6, line 41.	9 ---- 10-16
X ----- Y	ZEHNDER et al. Cross-linking Hybridization Assay for Direct Detection of Factor V Leiden Mutation. Clinical Chemistry. September 1997. Vol. 43. No.9., pages 1703-1708, see entire document.	9-12 ------ 13-16

NSDOCID: <WO___0047600A1_I_>

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 August 2000 (17.08.2000)

PCT

(10) International Publication Number
WO 00/47600 A1

(51) International Patent Classification[7]: C07H 21/00, 17/00, C12Q 1/68

(21) International Application Number: PCT/US00/03357

(22) International Filing Date: 10 February 2000 (10.02.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/119,417 10 February 1999 (10.02.1999) US
60/136,185 27 May 1999 (27.05.1999) US

(71) Applicant *(for all designated States except US)*: UNIVERSITY OF MARYLAND [US/US]; Office of Technology Commercialization, 6200 Baltimore Avenue, Suite 300, College Park, MD 20742-9520 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: AL-SHEIKHLY, Mohamad [US/US]; 1013 Aster Boulevard, Rockville, MD 20850 (US). BENTLEY, William, E. [US/US]; P.O. Box 253, St. Michaels, MD 20901 (US). SILVERMAN, Joseph [US/US]; 8101 Connecticut Avenue, S-407, Chevy Chase, MD 20815 (US).

(74) Agents: ESMOND, Robert, W. et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100 New York Avenue, N.W., Washington, DC 20005-3934 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(48) Date of publication of this corrected version:
13 September 2001

(15) Information about Correction:
see PCT Gazette No. 37/2001 of 13 September 2001, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/47600 A1

(54) Title: PHOTO-INDUCED NUCLEIC ACID HYBRIDIZATION

(57) Abstract: Photo-induced nucleic acid hybridization is achieved by exposure of a single-stranded nucleic acid molecules to ultraviolet (UV), visible (VIS) and near infrared (NIR) light. Specifically, irradiation increases the concentration of hydrogen bonded double-strand nucleic acid molecules as a result of complementary base pairing. Hybridization at pH 7.8 is most prevalent using UV (300 nm) irradiation, but is detectable even with NIR (920 nm) irradiation. The results offer promise of practical application in technologies related to genome arrays, northern and Southern blotting techniques, PCR, and hybrid nucleic acid-memory devices.

SDOCID: <WO_____0047600A1_IA>

Photo-Induced Nucleic Acid Hybridization

Field of the Invention

The invention is related to rapid nucleic acid hybridization by photo-induction at room temperature. Photo-induced nucleic acid hybridization can be used in the development of a nucleic acid memory chip for a hybrid system with massively parallel data searching capabilities wherein a directed light source induces hybridization reactions, and in nucleic acid amplification, e.g., via the polymerase chain reaction (PCR), and other techniques based on hybridization.

Background of the Invention

The fundamental chemistry of DNA is based on the double helix and the principal of complementarity. Hybridization is the hydrogen-bonding interaction between two DNA strands that obey the complementary rules. DNA strands can be very long, thread-like polynucleotides, made-up of a large number of deoxyribonucleotides containing purine and pyrimidine bases, which carry genetic information, and sugar and phosphate groups, which perform a structural role.

The variable part of DNA is its sequence of bases. DNA contains four kinds of bases, two purines and two pyrimidines. The two purines are adenine (A) and guanine (G). The two pyrimidines are thymine (T) and cytosine (C). Each strand of DNA is a covalently linked polymer, wherein each unit consists of a constant part (the sugar-phosphate "backbone") and one of either adenine, cytosine, guanine, or thymine. Each strand has a 3' and 5' end. When DNA forms a double-stranded helix, the strands must be anti-parallel (each end consists of a 3'-5' match) and complementary bases must align. Thus, a double-stranded helix will form through hydrogen bonding between the bases of two single strands. Hydrogen bonding occurs between A and T, and between G and C to produce complementary base pairs. This process is called annealing or hybridization. Ligation is the process in which the free 5' end of one strand reacts with the free 3' end of another strand to form a covalent bond to create a longer chain.

)NSDOCID: <WO___0047600A1_IA>

Although the structure of DNA helices is extremely rigid, it can be changed under certain conditions.

RNA, a single-stranded molecule, has a similar structure to DNA in terms of complementarity and polymeric structure. The variable part of RNA is also its sequence of bases. RNA contains four kinds of bases, two purines and two pyrimidines. The two purines are adenine (A) and guanine (G). The two pyrimidines are uracil (U) and cytosine (C). Each strand of RNA is a covalently linked polymer, wherein each unit consists of a constant part (the sugar-phosphate "backbone") and one of either adenine, cytosine, guanine, or uracil. Each strand has a 3' and 5' end.

The characteristics of nucleic acids are being manipulated in areas of technology such as DNA computing first described by Adleman, L.M., *Science 266:* 1021-1024 (1994), in an article entitled "Molecular computation of solutions to combinatorial problems." DNA computing addresses problems which require looking at several possible solutions (combinatorial problems) and solving those problems at the molecular level with nucleic acid molecules. There are many ongoing research efforts in DNA computing. *See, e.g.,* the extensive bibliography of molecular computing and splicing systems maintained by Pierluigi Frisco on the World Wide Web (*A Bibliography of Molecular Computation and Splicing Systems* (visited January 3, 2000) http://www.wi.leidenuniv.nl/~pier/dna.html). In addition, Baum has discussed a concept for an addressable memory system (Baum, E.B., *Science 268:* 583-585 (1995)). However, this memory system is in a flask and, as such, is not practical.

Among the ongoing efforts in DNA computing, the majority of the efforts are based upon computing in aqueous solutions. Among the principal activities in surface-based DNA computing are those being conducted at the University of Wisconsin, led by Max G. Lagally (Materials), Robert M. Corn (Chemistry) and Anne E. Condon (Computer Science). This Wisconsin group maintains a publications list of their own (Corn Group Publications (visited January 3, 2000) http://corninfo.chem.wisc.edu/publications.html). In a series of papers (Smith,

NSDOCID: <WO___0047600A1_IA>

L.M., *et al.*, *J. Comput. Biol. 5:* 255-267 (1998); Frutos, A.G., *et al.*, *Nucl. Acids Res. 25:* 4748-4757 (1997); Cai, W., *et al.*, "The Power of Surface-Based Computation," *Proceedings of the First Annual International Conference on Computational Molecular Biology (RECOMB '97)*, Santa Fe, New Mexico (1997), Thiel, A.J., *Analytical Chem. 69:* 4948-4956 (1997)), the Wisconsin group has developed the rationale and foundation for surface-based DNA computing. Smith *et al.*, *J. Comput. Biol. 5:* 255-267 (1998), have indicated that "[c]omplex combinatorial mixtures of DNA molecules are immobilized on a surface and subsets are tagged and enzymatically modified in repeated cycles of the 'DNA computation.' When the computation is complete, the sequence of the DNA remaining is determined, yielding the computational result."

Smith *et al.* indicate that performing DNA operations on surfaces introduces some new issues of both a fundamental and practical nature. For example, one must take into account the fact that negligible diffusion rates for static molecules affect hybridization kinetics, that the net negative charge on or near the surface may repel negatively charged complements in the working medium, and that probes and complementary molecules may be inaccessible due to possible steric hindrance.

In addition, Smith *et al.* note that the use of all four bases (A, C, G, and T) to encode DNA is impractical because of the thermal stability problems occurring with a different G/C content. Encoding with two bases has the disadvantage of requiring longer oligomers to store data, but has the advantage of permitting the G/C content to be kept relatively constant.

Smith *et al.* also note that when synthesizing longer nucleotide sequences, the yield will be reduced exponentially with polymer length. For example, if each step (nucleotide attachment) has an efficiency of 99.5%, then synthesis of a 100-mer will produce only 0.995^{100} = 60% yield of the correct sequence. The remainder will be impurities and side products.

Frutos *et al.*, *Nucl. Acids Res. 25:* 4748-4757 (1997), have addressed the problem of finding the best sequences for storing data in oligonucleotides. They

3NSDOCID: <WO__0047600A1_IA>

considered the stability of oligonucleotides based on the percentage G/C content and optimal nucleotide order to minimize the number of partial matches during hybridization reactions.

Thiel *et al.*, *Analytical Chem. 69:* 4948-4956 (1997), were able to distinguish between single- and double-stranded DNA regions on a gold surface using the Surface Plasmon Resonance (SPR) imaging detection technique. The SPR technique requires no label on the DNA, but fluorescently-labeled targets were used for verification of this approach.

Another line of ongoing research is the development of self-assembled DNA monolayers on surfaces (Boland, T. and Ratner, B.D., *Langmuir 10*:3845-3852 (1994); Park, D., *et al., Macromolecules 28:* 2595-2601 (1995); Pease, A.C., *et al., Proc. Natl. Acad. Sci. USA 91*:5022-5026 (1994); McGall, G., *et al., Proc. Natl. Acad. Sci. USA 93*:13555-13560 (1996)). Affymetrix, Inc. has developed a practical technique for "reading" the hybridized DNA pixels by stimulated fluorescence.

Clearly, hybridization plays in important role in new technologies, such as DNA computing as well as in technologies known in the art, such as PCR. Thus, a method of hybridization that would allow one to more efficiently and effectively carry out such technologies would be desirable. For example, in DNA computing, a compact DNA memory module would provide the levels of portability and durability that are needed for a practical hybrid system that are not now achievable.

Summary of the Invention

In the present invention, photo-induced hydrogen bonding between bases on single-stranded nucleic acid molecules is achieved not only by ultraviolet (UV), but also by visible (VIS) and near infrared (NIR) light. For example, while not wishing to be bound to any particular theory, two adjacent polynucleotides, which are otherwise stable in monomeric form in aqueous solutions, may be brought

together by the exclusion of an excited water molecule (in the case of NIR) or by cation charge displacement (in the case of UV) so that hydrogen bonding occurs. This bonding can be detected, e.g., via gel electrophoresis and detectable markers such as fluorophore-tagged markers.

Thus, the invention relates to a method of inducing rapid hybridization at room temperature of one or more nucleic acid molecules comprising applying a photo-excitation source to the one or more nucleic acid molecules to induce hybridization.

In particular, the invention relates to a method of inducing nucleic acid hybridization comprising:

(a) immobilizing a monolayer of single-stranded nucleic acid molecules in a combinatorial array of pixels to a substrate;

(b) contacting the substrate with a mobile phase containing one or more target single-stranded nucleic acid molecules at least partially complementary to the immobilized nucleic acid molecules; and

(c) applying a photo-excitation source to at least one pixel position of the combinatorial array to induce hybridization between the immobilized nucleic acid molecules and the one or more target nucleic acid molecules.

The invention further relates to a method of inducing hybridization of two or more nucleic acid molecules which are at least partially complementary, comprising applying a photo-excitation source to the two or more nucleic acid molecules, whereby hybridization is induced.

The invention further relates to a method for amplifying at least one double-stranded nucleic acid molecule, comprising:

(a) providing a first and second primer, wherein the first primer is complementary to a sequence at or near the 3'-terminus of the first strand of the nucleic acid molecule to be amplified and the second primer is complementary to a sequence at or near the 3'-terminus of the second strand of the nucleic acid molecule to be amplified;

(b) mixing the first and second primers, and the nucleic acid molecule to be amplified, with one or more polypeptides having nucleic acid polymerase activity, e.g., one or more DNA polymerases, one or more RNA polymerases, one or more reverse transcriptases, etc., which may or may not be thermostable, thereby forming a mixture;

(c) illuminating the mixture with light under conditions favoring the photo-induced hybridization of the first primer to the first strand of the nucleic acid molecule, and the second primer to the second strand of the nucleic acid molecule;

(d) producing a third nucleic acid molecule complementary to the first strand and a fourth nucleic acid molecule complementary to the second strand;

(e) denaturing the first and third strands, and the second and fourth strands; and

(f) repeating steps (a) to (e) one or more times.

The invention further relates to a method of sequencing a nucleic acid molecule, comprising:

(a) mixing a primer with a first nucleic acid molecule to form a first mixture;

(b) illuminating the first mixture with a photo-excitation source under conditions favoring the photo-induced hybridization of the primer with the first nucleic acid molecule;

(c) mixing the first mixture with one or more deoxyribonucleoside triphosphates (dNTPs), one or more polypeptides having nucleic acid polymerase activity, e.g., one or more DNA polymerases, one or more RNA polymerases, one or more reverse transcriptases, etc., which may or may not be thermostable, and one or more terminator nucleotides, e.g., ddTTP, ddATP, ddGTP, ddITP or ddCTP, thereby forming a second mixture;

(d) incubating the second mixture under conditions sufficient to synthesize a random population of DNA molecules complementary to the first nucleic acid molecule, wherein the synthesized nucleic acid molecules are shorter

in length than the first nucleic acid molecule and wherein the synthesized nucleic acid molecules comprise a terminator nucleotide at their 3' termini; and

(e) separating the synthesized nucleic acid molecules by size so that at least a part of the nucleotide sequence of said first nucleic acid molecule can be determined.

The invention further relates to a nucleic acid memory device comprising a photo-excitation source, a substrate with nucleic acid molecules attached in a combinatorial array, a mobile phase containing nucleic acid target molecules at least partially complementary to the attached nucleic acid molecules, and a detector which detects hybridization within the combinatorial array.

Brief Description of the Figures

Fig. 1: A schematic diagram of a DNA memory module with an external photo-excitation source.

Fig. 2: A schematic diagram of a DNA memory module integrated with a semiconductor laser array.

Fig. 3A: A digital micromirror device that has two states - a "0" state wherein the mirror is tilted at +10 degrees and a "1" state wherein the mirror is tilted at -10 degrees. Fig. 3B: A schematic diagram of a DNA memory system using a DMD spatial light modulator.

Fig. 4: A DNA memory system employing a transmissive type SLM as the photo-excitation source.

Fig. 5: An illustration of storage and retrieval operations for DNA memory by sequencing.

Figs. 6A-6C: Photographs of hybridized DNA strands in aqueous solutions analyzed by gel electrophoresis. At each wavelength, an exposed sample, the same sample boiled, and the non-exposed control are shown.

Figs. 7A-7C: Photographs of hybridized DNA strands in aqueous solutions analyzed by gel electrophoresis.

3NSDOCID: <WO___0047600A1_IA>

Fig 8: A typical plot of the sum of the intensities of two bands plotted against the weight of an oligonucleotide.

Fig 9A: A graph showing an increase in absorption (A_{260}) of DNA samples with respect to temperature. Fig. 9B: A photograph of melting depicted by gel electrophoresis.

Fig 10 A graph showing the relative hybridization ratios of 37 μM oligonucleotide at pH 7.8 by photolysis in various regions: UV (300 nm), visible (450 nm) and NIR (920 nm) and the relative hybridization ratios of 37 μM oligonucleotide at various pH values. Note that the hybridization ratio (HR) was calculated as the intensity of light in gel (upper/lower) sample/ (upper/lower) unphotolyzed

Fig 11 A graph showing the effect on HR of the exposure time with 300 nm light on the hybridization yield of 37 µM at pH 7.8. Hybridization yield increases to a maximum value at 0.1 seconds. At higher exposure times (0.6 and 4.0 seconds), UV photolysis appears to induce scission of the hybridized molecules.

Fig 12 A graph showing the effect on the HR of a DNA sample (pH= 7.8) photolyzed in UV, visible and IR regions.

Fig. 13 A graph showing the results of a DNA melting experiment. "A" represents absorbance at 260 nm which increases as temperature increases demonstrating that the double-stranded oligonucleotide is dissociated into two UV-light absorbing monomers. "B" represents the scanning density of the upper band in agarose gel which decreases demonstrating dissociation into two UV-absorbing monomers A corresponding increase in the lower band was also observed (not shown)

Fig. 14 A diagram of the nucleotide sequence (SEQ ID NO:1) used in Example 1.

Detailed Description of the Preferred Embodiments

Typically, manipulation of DNA hybridization is made via changes in temperature. Higher temperatures cause melting of double-stranded DNA, while lower temperatures affect annealing, which are critical steps in PCR (U.S. Patent Nos. 4,683,195 and 4,683,202) and other technologies. Hybridization is also affected by buffer ionic strength and the use of detergents. The possibility that light can be used as a further control variable to enhance or otherwise tailor the annealing step has an enormous impact on the design and use of hybridization-based technologies such as techniques using genome arrays and northern and Southern blotting analytical techniques. Further, photo-induced hybridization substantially improves techniques based on enzyme extensions of nucleic acid hybridization such as PCR. Additionally, the realization of DNA-based computing (Adleman, L. M., *Science 266*:1021-1024 (1994); Frutos, A.G., *JACS 120*:10277-10282 (1998)) is greatly facilitated if a photochemical reaction can be found to induce hybridization of nucleic acid molecules efficiently since rapid photo-induced hybridization can provide a more realistic basis for nucleic acid-based information storage than slow enzyme-based chemistry. Thus, photonic interfaces (as shown in Fig. 2) would provide a degree of spatial control of nucleic acid hybridization which is unavailable by other means. In Fig. 2, the photonic interface is between the VCSEL array of the NIR emitting lasers and the target DNA gel.

Rapid hybridization at room temperature, in the manner described here, is useful for genomic chip-based technology, nucleic acid amplification and sequencing, and other laboratory techniques based on nucleic acid hybridization. The standard PCR technique, disclosed in U.S. Patent Nos. 4,683,195 and 4,683,202, is a technique that has been used for amplifying and sequencing nucleic acid molecules. In this technique, a sample containing the nucleic acid sequence to be amplified or sequenced (the "target sequence") is first heated to denature or separate the two strands of the nucleic acid. The sample is then cooled and mixed

3NSDOCID: <WO 0047600A1 IA>

with specific oligonucleotide primers which hybridize to the target sequence. Following this hybridization, for amplification of the target sequence, DNA polymerase in a buffered aqueous solution is added to the sample, along with a mixture of the dNTPs that are linked by the polymerase to the replicating nucleic acid strand. After allowing polymerization to proceed to completion, the products are again heat-denatured, subjected to another round of primer hybridization and polymerase replication, and this process repeated any number of times. Since each nucleic acid product of a given cycle of this process serves as a template for the production of two new nucleic acid molecules (one from each parent strand), the PCR process results in an exponential increase in the concentration of the target sequence. Thus, in a well-controlled, high-fidelity PCR process, as few as 20 cycles can result in an over one million-fold amplification of the target nucleic acid sequence.

By incorporating photo-induced hybridization in methods of nucleic acid amplification and sequencing, the annealing step performed after cooling at a reduced temperature may be eliminated. Instead, one may do the annealing and elongating at the same temperature. Additionally, one may tailor the specificity of various primers hybridized to the nucleic acid based on light excitation. That is, the incident light affords another control variable for fine tuning the amplification and sequencing reactions. This is an enormous advantage over the current state-of-the-art.

Other methods of nucleic acid amplification and sequencing may analogously employ the photo-induced hybridization methods of the invention. Examples of such amplification techniques include, but are not limited to, Strand Displacement Amplification (SDA; U.S. Patent No. 5,455,166; EP 0,684,315), and Nucleic Acid Sequence-Based Amplification (NASBA; U.S. Patent No. 5,409,818; EP 0 329 822). Examples of sequencing techniques include, but are not limited to, PCR-based nucleic acid fingerprinting techniques such as Random Amplified Polymorphic DNA (RAPD) analysis (Williams, J.G.K., *et al.*, *Nucl. Acids Res. 18(22):* 6531-6535 (1990)), Arbitrarily Primed PCR (AP-PCR; Welsh,

J , and McClelland, M., *Nucl. Acids Res. 18(24):* 7213-7218 (1990)), DNA Amplification Fingerprinting (DAF: Caetano-Anollés *et al.*, *Bio/Technology 9:* 553-557 (1991)), microsatellite PCR or Directed Amplification of Minisatellite-region DNA (DAMD; Heath, D.D., *et al.*, *Nucl. Acids Res. 21(24):* 5782-5785 (1993)), and Amplification Fragment Length Polymorphism (AFLP) analysis (EP 0 534 858, Vos, P., *et al.*, *Nucl. Acids Res. 23(21):* 4407-4414 (1995); Lin, J.J., and Kuo, J , *FOCUS 17(2):* 66-70 (1995)). In particular, the methods of the present invention will be useful in the fields of medical therapeutics and diagnostics, forensics, and agricultural and other biological sciences, in any procedure using hybridization of nucleic acid molecules.

Thus, in one embodiment, the invention relates to a method for amplifying at least one double-stranded nucleic acid molecule, comprising:

(a) providing a first and second primer, wherein the first primer is complementary to a sequence at or near the 3'-terminus of the first strand of the nucleic acid molecule to be amplified and the second primer is complementary to a sequence at or near the 3'-terminus of the second strand of the nucleic acid molecule to be amplified;

(b) mixing the first and second primers, and the nucleic acid molecule to be amplified, with one or more polypeptides having nucleic acid polymerase activity, e.g., one or more DNA polymerases, one or more RNA polymerases, one or more reverse transcriptases, etc., which may or may not be thermostable, thereby forming a mixture;

(c) illuminating the mixture with light under conditions favoring the photo-induced hybridization of the first primer to the first strand of the nucleic acid molecule, and the second primer to the second strand of the nucleic acid molecule;

(d) producing a third nucleic acid molecule complementary to the first strand and a fourth nucleic acid molecule complementary to the second strand;

(e) denaturing the first and third strands, and the second and fourth strands; and

3NSDOCID: <WO___0047600A1_IA>

(f) repeating steps (a) to (e) one or more times.

In preferred embodiments, the denaturing and hybridizing steps are performed at temperatures below about 90°C, below about 85°C, below about 80°C, below about 75°C, below about 70°C, below about 65°C, below about 60°C, below about 55°C, below about 50°C, below about 45°C, or below about 40°C, or at about 20°C to about 45°C.

In a related embodiment, the invention also provides a method of sequencing a nucleic acid molecule, comprising:

(a) mixing a primer with a first nucleic acid molecule to form a first mixture,

(b) illuminating the first mixture with a photo-excitation source under conditions favoring the photo-induced hybridization of the primer with the first nucleic acid molecule;

(c) mixing the first mixture with one or more deoxyribonucleoside triphosphates (dNTPs), one or more polypeptides having nucleic acid polymerase activity, e.g., one or more DNA polymerases, one or more RNA polymerases, one or more reverse transcriptases, etc., which may or may not be thermostable, and one or more terminator nucleotides, e.g., ddTTP, ddATP, ddGTP, ddITP or ddCTP, thereby forming a second mixture;

(d) incubating the second mixture under conditions sufficient to synthesize a random population of DNA molecules complementary to the first nucleic acid molecule, wherein the synthesized nucleic acid molecules are shorter in length than the first nucleic acid molecule and wherein the synthesized nucleic acid molecules comprise a terminator nucleotide at their 3' termini; and

(e) separating the synthesized nucleic acid molecules by size so that at least a part of the nucleotide sequence of said first nucleic acid molecule can be determined.

In a related aspect, the primer, first nucleic acid molecule, one or more dNTPs, one or more polypeptides having nucleic acid polymerase activity, and one or more terminator nucleotides may be mixed together prior to illuminating the

mixture with the photo-excitation source under conditions favoring the photo-induced hybridization of the primer with the first nucleic acid molecule, such that hybridization and synthesis of the random population of nucleic acid molecules may occur contemporaneously.

In addition to nucleic acid amplification and sequencing and other laboratory techniques, photo-induced hybridization can be used to obtain a high-density hybrid nucleic acid memory and massively parallel data searching system in an integrated, compact package, nucleic acid molecules, vectors containing such molecules, host cells containing such molecules, etc. Massively parallel searching is enabled by simultaneous hybridization reactions of all nucleic acid complementary pairs as described below.

Thus, another aspect of the invention relates to a method of inducing nucleic acid hybridization comprising:

(a) immobilizing a monolayer of single-stranded nucleic acid molecules in a combinatorial array of pixels to a substrate;

(b) contacting the substrate with a mobile phase containing one or more target single-stranded nucleic acid molecules at least partially complementary to the immobilized nucleic acid molecules; and

(c) applying a photo-excitation source to at least one pixel of the combinatorial array to induce hybridization between the immobilized nucleic acid molecules and the one or more target single-stranded nucleic acid molecules.

Another aspect of the invention relates to a method of inducing nucleic acid hybridization comprising:

(a) obtaining a substrate comprising a monolayer of single-stranded nucleic acid molecules in a combinatorial array of pixels immobilized thereon and a mobile phase containing one or more target single-stranded nucleic acid molecules at least partially complementary to the immobilized nucleic acid molecules; and

(b) applying a photo-excitation source to at least one pixel of the combinatorial array to induce hybridization between the immobilized nucleic acid molecules and the one or more target single-stranded nucleic acid molecules.

Photo-induced hybridization, according to this aspect of the invention, involves certain components in contact with, and layered upon, a semiconductor or other substrate. The components may include a photo-excitation source and its associated opto-electronics, a self-assembly of a monolayer of nucleic acid molecules immobilized on the substrate surface, and a mobile phase, e.g., a buffered salt solution, a gel matrix, etc., carrying solubilized nucleic acid at least partially complementary to the immobilized nucleic acid molecules. The nucleic acid molecules may be chemically affixed directly to the photo-excitation array or to an intermediate highly transmittive insulating layer such as fused silica.

The photo-excitation source may be internal or external to the substrate. As an example of an internal source, the substrate may comprise one more light-producing or emitting components that can be induced, e.g., electronically, electrically or physically, to produce light at a localized site, thereby hybridizing the nucleic acid molecules immobilized at that site. As an example of an external source, the substrate containing the immobilized nucleic acid molecules may be illuminated with a light source, thereby hybridizing the nucleic acid molecules. Detailed examples of internal and external light sources are provided below.

In one embodiment of the nucleic acid memory and searching system, the system is a nucleic acid memory device in the form of a computer chip interfaced with an electronic computer. Such chips, containing a potentially enormous database of hybridized nucleic acid molecules, with their near real-time massive parallel searching capabilities, can be applied to areas of target identification, parts location, and personnel identification, e.g., fingerprint and other bio-source identification, multi-level security, low power robotics with persistent memory (for powering down), "smart dust," i.e., memory units for nanoscale sensor arrays, "smart drugs," i.e., drugs that are able to respond to "invisible" conditions, among

NSDOCID: <WO___0047600A1_IA>

others. Additionally, photo-induced hybridization could be used to as a diagnostic technique in drug design, among other applications.

Photo-Induced Hybridization Device

A device according to the present invention may be used as a high-density hybrid nucleic acid memory and massively parallel data search system. The compact device may be assembled in a layered sandwich configuration comprising, in any order: a photo-excitation source (which is also possibly the substrate or is external to the device), a substrate with nucleic acid molecules immobilized thereon in a combinatorial array (pixels), a mobile phase containing target nucleic acid molecules complementary to the immobilized nucleic acid molecules, and a detector which senses hybridization at each pixel once the hybridization process is complete. The location of the detector may vary as shown in Figs. 1 and 2 which illustrate particular embodiments of the device. An electronic computer may be connected to the device to facilitate the data searching capability.

Substrate Layer and Internal Photo-Excitation Source

In particular, the substrate layer comprises well-defined, self-assembled nucleic acid molecules such as a monolayer of identical single-stranded nucleic acid molecules in a combinatorial array. This means that distinct pixel areas, i.e., areas containing particular nucleic acid molecules, on the substrate contain different nucleic acid sequences and that each pixel can be addressed individually. The self-assembly of the monolayer of nucleic acid molecules on the substrate surface must be accomplished with sufficient surface concentration and stability. This can be accomplished by placing a minimum of about 10^5 molecules/pixel (based on currently available pixel size well-known to one of ordinary skill in the art) and conducting the hybridization at room temperature.

3NSDOCID: <WO___0047600A1_IA>

Cai *et al.*, "The Power of Surface-Based Computation," *Proceedings of the First Annual International Conference on Computational Molecular Biology (RECOMB '97)*, Santa Fe, New Mexico (1997), note that good attachment chemistry ensures that the properly prepared nucleic acid molecules are immobilized to the substrate surface at a high density, and that other nucleic acid molecules exposed to the substrate surface later, for example, during hybridization, do not bond non-specifically to the substrate surface. Preliminary experiments on glass surfaces indicate that different chemistry can be used to attach either the 3' or 5' end of an oligonucleotide to the surface. However, one problem with glass is the non-specific binding of DNA molecules to the surface, i.e., the molecules bind to the surface itself rather than to their complement. Thus, thermally grown oxides on silicon wafers or alkanethiol self-assembled monolayers on gold surfaces may be better alternatives.

The substrate to which the monolayer is attached may be a photo-excitation source such as an integrated monolithic array of opto-electronic devices developed to assure a collimated monochromatic light source for transmission near the optimal wavelength which effectively induces nucleic acid hybridization, i.e., a wavelength between about 350 nm and about 1200 nm. The monochromatic light is the source of energy for the hybridization reactions between the nucleic acid molecules and the surrounding mobile phase medium. Preferred substrates include semiconductors, i.e., vertical cavity surface emitting laser (VCSEL) arrays. More preferably, the photo-excitation source includes semiconductor wafers such as GaAs or GaN. Compound semiconductors, such as GaAs, an excellent source of NIR light (850-950 nm), provide the activation energy for the hybridization reactions between the immobilized nucleic acid molecules and the surrounding target nucleic acid molecules. Towe *et al.* (Ramos, P.A. and Towe, E., *Appl. Phys. Lett. 69*:3321-3323 (1996); Towe, E. and Ramos, P.A., "Active (In, Ga) As/GaAs Blue-Green Light-Emitters," SPIE Photonics West, Optical Interconnects Symposium, San Jose, CA, Feb. 22, 1997; Ramos, P.A. and Towe, E., *Appl. Phys. Lett. 68:* 1754-1756 (1996); Ramos, P.A. and Towe, E., *Optics*

NSDOCID: <WO___0047600A1_IA>

Comm. *132*:121-127 (1996)) have shown that 112 GaAs wafers with [112]-oriented (In,Ga)As/GaAs epitaxial layers can produce coherent blue radiation (the frequency of the second harmonic). The semiconductor may also be used to detect differences in the electrical permittivity of the hybridized and unaltered pixels as described below.

External Photo-Excitation Source

As noted above, the energy required to induce hybridization may alternatively be provided by an external light source, such as a laser. It is necessary to determine the intensity and duration of the laser sufficient to hybridize enough immobilized molecules to be detectable. In order to identify the optimum wavelength, tunable lasers covering from about 260 nm in the UV range to about 10 nm in the NIR range, for example, may be used. An exemplary tunable laser can produce light pulses as short as 100 femtoseconds with peak power as high as several terawatts. Also, continuous sources of light with wavelengths in the optimum region may be used in PCR and blotting applications.

To efficiently deliver the light source for photo-induced hybridization, spatial light modulation (SLM) techniques may be applied to construct a programmable pattern generator that can be electronically programmed to illuminate specific areas on the surface so as to activate the hybridization process. Within the photo-excitation system, the SLM device controls the external light to address different pixels on the immobilized nucleic acid array. The SLM device can turn the light on and off at each pixel independently. There are two types of SLM devices: reflective type and transmissive type. Each type of SLM device has its own advantages and selection depends on the working wavelength. In either case, the selected SLM device may be controlled by a computer that also reads the data collected from the nucleic acid memory system.

One particular reflective type SLM device that may be used to photo-induce nucleic acid hybridization is the digital micro-mirror device (DMD)

)NSDOCID: <WO___0047600A1_IA>

developed by Texas Instruments. The DMD is a reflective type device containing a plurality of mirrors that functions well in a broad spectral range from UV to NIR. Fig. 3 shows a schematic drawing of such a photo-induced DNA hybridization device. The DMD mirrors have two states when applying voltage onto them: "0" state at which the mirrors are tilted +10° and "1" state at which the mirrors are titled -10°. The incident light is directed to the light trap at the "0" state while the light goes to a lenslet array at the "1" state. The "0" and "1" states at each micromirror are determined by the mask pattern needed to correctly excite the nucleic acid hybridization process. The mask pattern chosen prevents the photo-excitation source from reaching particular nucleic acid molecules such that only desired immobilized nucleic acid molecules are subjected to the photo-excitation source. Masking is well-known in the art as illustrated by Pirrung, M.C., *Chem. Rev. 97:* 473-488 (1997) and U.S. Patent No. 5,744,101. The lenslet array directs the light beam on the semiconductor substrate, for example, to selectively induce the reactions. This unique feature allows the use of the optimum light source for photo-induced nucleic acid hybridization.

Transmissive-type SLM devices include a liquid crystal display and can be integrated with a nucleic acid memory device directly. Thus, the system may be more compact and simple in construction. However, most common electro-optical materials, such as liquid crystals, absorb UV light thus limiting their usage in the UV light wavelength region. Should the optimum wavelength for nucleic acid hybridization be in this wavelength range, then special developments are required to construct devices that are transparent in UV light wavelength range. PLZT, a solid-state material, is a significant potential candidate for an SLM device operating at UV wavelengths. *See* Fig. 4.

There are at least three methods of interfacing optics with aqueous systems - free space optical delivery system, guided wave optical system, and a combination of the two. Once the optimum photo-excitation conditions are found, the integration of a nucleic acid memory system with potential laser systems can be facilitated to provide the highest yield of photo-induced hybridization.

Mobile Phase

The mobile phase may comprise an aqueous solution, e.g., a buffered salt solution or a solid or semi-solid phase such as a gel matrix, containing solubilized nucleic acid target molecules that are at least partial complements to the immobilized nucleic acid molecules. The mobile phase can be in direct contact with the photo-excitation source or immobilized array by solubilizing the target nucleic acid molecules into an agarose or a polyacrylamide gel or a gel containing PAA or PEO using methods well-known in the art. The mobile phase preferably comprises either at least partially complementary strands to all the immobilized nucleic acid sequences or long nucleic acid molecules to be stored by sequencing methods. However, the target strands have to be at least partially complementary for hybridization to occur. The target nucleic acid molecules in the mobile phase may have well-defined detectable labels, e.g., fluorescent, radioactive, phosphorescent, bioluminescent, chemiluminescent, biochemical or other labels that permit detection of the target nucleic acid molecules in the mobile phase or upon the substrate after hybridization, and should be stable over a broad range of thermodynamic and ambient light conditions. In addition, the target nucleic acid molecules should have sufficient mobility to diffuse to the immobilized nucleic acid molecules for hybridization to occur.

In order to create the most favorable conditions for the hybridization reaction with the immobilized nucleic acid molecules, the nucleic acid molecule concentration in the mobile phase should be maximized. Thus, the effects of pH, temperature, additives such as buffer and enzyme, and water content on the nucleic acid molecule concentration in the mobile phase should be determined as described below.

Nucleic Acid Selection

Nucleic acid molecules such as DNA, cDNA, RNA, mRNA and tRNA of at least eight nucleotides in length are useful in the invention. When UV light is employed, preferred nucleic acid molecules do not contain thymine or contain a small percentage of thymine. Additionally, nucleic acid molecules may contain a mismatch at every third base. Exemplary nucleic acid molecules is set forth in SEQ ID NOS:1 and 2.

Other nucleic acid molecules useful in the invention include nucleic acid molecules that have a melting temperature below the operating temperature of the photo-induced hybridization device. Methods of determining melting temperatures are well-known in the art.

Hybridization Detectors

Hybridization reactions between the immobilized nucleic acid molecules and the target nucleic acid molecules may be induced by light emitted at addressable locations. Information may be stored at specific addresses in the form of hybridized (double-stranded) polynucleotides. Retrieval of the data at each pixel may be made possible by means of distinguishing single- from double-stranded molecules. Possible methods for this purpose include tagging the target nucleic acid molecules with radioactive, fluorescent, phosphorescent, bioluminescent or chemiluminescent moieties, or another type of label which can be stimulated by an external light source and detected by, for example, a confocal detector or a charged coupled device (CCD) chip above the substrate. In either case, the object is to detect hybridized strands by a signal induced by illuminating the substrate with another light source. The confocal detector scans the substrate and collects data to form an image of the array. The CCD chip similarly provides an image of the array. When fluorescent labels are used, the fluorochrome on the single-stranded nucleic acid molecule preferably emits light at a different

NSDOCID: <WO___0047600A1_IA>

wavelength and a lower fluorescence yield, than that on the hybridized double-stranded nucleic acid molecule.

Light can illuminate the target nucleic acid molecules and be collected effectively by the detector array due to the following: (I) the immobilized layer is a layer of immobilized single-stranded nucleic acid molecules. If the nucleic acid molecules are DNA, for example, then the lateral dimension of the nucleic acid molecules is about 9Å and the space between them is about 20Å. Therefore, most of the light can go through without being affected; (ii) the small dimension determines that no diffraction will occur; and (iii) the thickness of both the immobilized layer and the target layer are small and the total energy being absorbed is relatively small.

To facilitate the detection of hybridization on the substrate, the single-stranded nucleic acid molecules may be labeled with, for example, radioactive isotopes, fluorescent labels, phosphorescent labels, chemiluminescent labels, bioluminescent labels and enzyme labels. U.S. Patent No. 4,581,333 describes the use of enzyme labels to increase sensitivity in a detection assay. Radioisotopic labels are disclosed in U.S. Patent Nos. 4,358,535, and 4,446,237. Fluorescent labels (EP 144,914), chemical labels (U.S. Patent Nos. 4,582,789 and 4,563,417) and modified bases (EP 119,448) may also be employed. Fluorescent labels are most preferred. Specific fluorescent labels include digoxigenin via 3' end labeling with DIG-dUTP and fluorescein-12-dUTP or 6-FAM phosphoramidite. In the course of the analysis, the fluorochrome-labeled hybridized strands are excited by a light source such as laser and the emitted fluorescent light is detected to generate an image of the bands in the gel (Adleman, L.M., *Science 266*: 1021-1024 (1994)).

Labels may be employed prior to solubilization of the target nucleic acid molecules in the mobile phase. Alternatively, the immobilized nucleic acid molecules may be labeled. The nucleic acid molecules are labeled with at least one detectable label internally and/or at or near the 3' and/or 5' termini. Multiple labels, which may be the same or different, may be used.

NSDOCID: <WO__0047600A1_IA>

As an alternative to labeling the nucleic acid molecules, the substrate, e.g., semiconductor, can serve as the light source for post-hybridization absorption studies, e.g., as a means of distinguishing hybridized nucleic acid molecules from unaltered nucleic acid molecules in the detection process.

Several methods have been suggested in the literature for detecting hybridization on chips (Smith, L.M., *et al.*, *J. Comput. Biol. 5:* 255-267 (1998)). These include electrical, linear and non-linear optical, and fluorescence methods. Specifically, options include Fourier transform infrared (FTIR) spectroscopy, surface plasmon resonance (SPR), and diffraction and spectroscopy methods.

Following the hybridization process, the nucleic acid fragments may be analyzed by gel electrophoresis wherein the fragments are separated according to their size (molecular weight). The original sequence can be read from the separated bands. Electrophoresis of nucleic acid molecules may be carried out in an agarose or polyacrylamide gel.

Detecting hybridization requires one to distinguish between hybridization on the substrate surface from hybridization in the mobile phase or from hybridization with neighboring pixels, and, second, to determine in which pixels hybridization has occurred. The success of this second step requires identifying the same pixels that were lased to within tolerable error limits. Errors may arise due to several factors. For example, there may be partial matching hybridization of target nucleic acid molecules in the mobile phase and diffusion of these molecules away from the pixel locations. There may also be hybridization at adjacent pixels due to poor laser registration or focusing, or from conduction/convection heating from the hybridization reactions in the designated pixel(s). Thus, the substrate may be prepared by exciting selected pixel locations with an external source such as, for example, laser light, to initiate the hybridization reactions in that area.

Verification of hybridization can be determined by performing experiments wherein the gel is stripped from the substrate so that the skilled artisan can work directly with the substrate surface nucleic acid molecules. Fluorescence and electrophoresis are reliable and standardized techniques that may be applied here.

\SDOCID: <WO___0047600A1_IA>

The fluorescence experiments qualify the extent to which neighboring nucleic acid molecules are affected. Electrophoresis identifies the products of the hybridization, such as, for example, nucleic acid molecule experiment extensions by ligation reactions. Such experiments entail chemically or mechanically removing the nucleic acid molecules from the substrate.

Having shown that nucleic acid molecules on the substrate are hybridized under the action of laser light, the next task is to determine the hybridized nucleic acid molecule locations. As indicated above, several methods are known for detecting hybridization on substrates such as chips, including electrical, linear and non-linear optical, and fluorescence methods. Specifically, options include Fourier transform infrared (FTIR) spectroscopy, surface plasmon resonance (SPR), and diffraction and spectroscopy methods.

Hybridization Reactions

To induce hybridization, the light source, such as the VCSEL array, is focused onto each selected pixel of the immobilized nucleic acid array by a beam condensor/conditioner array such as a diffractive lenslet array. The photo-induced hybridization reactions take place on the immobilized array. The lenslet array may be fabricated from semiconductor materials such as, for example, glass, quartz, etc., depending upon the wavelength of the VCSEL array. For the UV range, preferred hybridization conditions include irradiation at about 300 nm with an intensity of about 14 mW/cm^2 from about 0.1 seconds to about 4.0 seconds. For the visible range, the preferred hybridization conditions include irradiation at about 430 to about 514 nm with an intensity of about 1.0 mW/cm^2 from about 0.4 seconds to about 30 seconds. For the NIR range, the preferred hybridization conditions include irradiation at about 920 nm with an intensity of about 4.45 mW/cm^2 from about 0.2 seconds to about 15 seconds.

3NSDOCID: <WO___0047600A1_IA>

The photo-excitation source initiates the following hybridization reactions:

nucleic acid (mobile phase) + nucleic acid (immobilized) = hybridized nucleic acid (1)

nucleic acid (mobile phase) + nucleic acid (mobile phase) = hybridized nucleic acid (2)

While the products of the reaction (1) will remain attached to the substrate surface, the products of reaction (2) will be present in the mobile phase. In addition, it is expected that, under certain conditions, photolysis will also lead to the formation of cyclobutane thymine dimers in the gel (discussed below). The nucleic acid molecules attached to the substrate are at the center of the proposed hybrid system, therefore reaction (1) is the focus of the invention. The product of reaction (2) is also hybridized nucleic acid molecules, but is present in the mobile phase instead of on the surface. These nucleic acid molecules may simply constitute a noise level in the system

The complications that are raised due to reaction (2) and the dimerization reactions can be solved by vigorously washing the products of these reactions. Since these products are hosted in the mobile phase, aqueous-based solutions are excellent solvents that may be used for the washing process. Another technique to reduce the complications of reaction (2) and the dimerization reactions is the application of an external electrical field on the mobile phase since the nucleic acid molecules are negatively charged. This will force the hybridized nucleic acid in the mobile phase to migrate to the anode area.

Dimerization and Photo-Induced Hybridization

One of the objectives of the claimed invention is to achieve nucleic acid hybridization and crosslinking reactions of complementary strands while simultaneously minimizing dimerization reactions. UV-induced ligation and UV-induced covalent dimerization of oligonucleotides are well-known (Setlow, J. K., "The Molecular Basis of Biological Effects of Ultraviolet Radiation and Photoreactivation," In *Current Topics in Radiation Research*, Ebert M. & Howard,

A. (eds.), North-Holland Publishing Co., Amsterdam 2:195-248 (1966); Rahn, R. O. and Patrick, M. H. "Photochemistry of DNA, secondary structure, photosensitization, base substitution, and exogenous molecules," in *Photochemistry and Photobiology of Nucleic Acids*, Wang S.Y. (ed.) Academic Press, New York, Biology 2:97-145 (1976); Carell, T., *Angew. Chem. Int. Ed. Engl. 34*:2491-2494 (1995); Love, J. D., *et al., J. Biol. Chem. 261*:10051-10057 (1986); Kypr, J., *et al., J. Biomol. Struct. Dyn. 11*:1225-1236 (1994)). UV light has been applied in DNA chemistry in a variety of applications despite the belief that its major effect is the creation of cyclobutane thymine dimers (Marmur, J., and Grossman, L., *Proc. Natl. Acad. Sci. U SA 47*:778-787 (1961); Rahn, R., *et al., Biophys. J. 9*:510-517 (1969), Glisin, V.R., and Doty, P., *Biochim. Biophys. Acta 142*:314-322 (1967); Patrick, M.H., *Photochem. Photobiol. 25*:357-372 (1977)).

Recently, Pospíšilová and Kypr (Pospíšilová, S. and Kypr, J., *Photochem. Photobiol. 65*:945-948 (1997)) have demonstrated UV light-induced crosslinking of DNA complementary strands. They found that the yield of crosslinked DNA was dependent on the DNA strand length and the UV light dose. Tiedge (Tiedge, H., *DNA Cell Biol. 10*:143-147 (1991)) has used UV light-induced cross-linking to immobilize target RNA molecules for *in situ* hybridization. Similarly, Affymetrix, Inc. (McGall, G.H., *et al., J. Am. Chem. Soc. 119*:5081 (1997); Anderson, R.C., *et al., Topics in Curr. Chem. 194*:117-129 (1998)) has used near-UV light for the synthesis of polynucleotide probe arrays on surfaces. In each of these applications, the results would have been hindered if dimerization were the controlling reaction, but the results were not hindered.

The issue of dimerization has also been addressed by Carell (Carrell, T., "Sunlight-Damaged DNA Repaired with Sunlight," *VCH Verlag GmbH*, Weinheim, *Angew Chem. Int. Ed. 34(22)*: 2491 (1995)). Carell has shown that the yield of photo-induced dimers depends on both the structure and sequence of the DNA as well as on the pH of the aqueous media. In the reversible reaction, a *cis-trans* cyclobutane thymine dimer forms upon UV light radiation (from about 200 to about 300 nm) of DNA.

BNSDOCID: <WO 0047600A1 IA>

Several steps can be taken to reduce the probability of formation of cyclobutane thymine dimers:

1. The backward reaction can be enhanced by using another radiation source in the range of from about 300 to about 500 nm.

2. The concentration of thymine can be reduced by selecting the sequences of nucleic acid molecules. Since the dimerization reaction is strongly dependent on the structure and sequence of the nucleic acid molecules, one may use single-stranded nucleic acid molecules with a little or no thymine. In this case, the hybridization reactions will depend more strongly on the guanine-cytosine pair (with three hydrogen bonds) than on the adenine-thymine pair (with two hydrogen bonds). Double helices with guanine and cytosine are more stable than those with adenine and thymine owing to the stronger hydrogen bonding.

3. Various quenchers can be used to suppress photo-induced dimerization. It has been demonstrated that dimerization is virtually suppressed when the pyrimidine oligonucleotides $d(TC)_y$ or $d(C)_m$, are added to DNA carrying $d(TC)_x$ or $d(C)_n$ containing inserts, respectively (Lamichev, V.I., *et al.*, *Nature 344*:568-570 (1990)).

4. Photo-induced dimerization is also pH-dependent owing to deprotonation reactions. Thus, the pH factor will be used to further suppress the dimerization reactions.

Characterization of Photo-Induced Hybridization Mechanisms and Kinetics

Characterization and optimization of photo-induced hybridization reactions require determination of the wavelength, intensity, pulse length, and pulse repetition frequency to maximize the yield of hybridized nucleic acid molecules. Such determinations may be made using mobiles phases such as aqueous solutions and gels and an external tunable laser or in semiconductor chips using an internal excitation source. Analysis of the results may be conducted using classical

fluorescence and electrophoresis experiments as well as advanced techniques for surface interrogation.

While not wishing to be bound by any particular theory, it is believed that the mechanism of the photo-induced hybridization reactions is based on photo-excitation that leads to the hydrogen bonding of the bases of two nucleic acid molecules as described above. The flash photolysis technique (described below) may be used to determine kinetic measurements, identify the chemical structure of the nucleic acid molecule in the excited state, and elucidate the reaction mechanism of the nucleic acid molecule and its moieties that can be applied in computational processes

Flash photolysis is a powerful technique for studying transient species (mainly excited states and radicals) produced by photo-induced excitation of molecules using short pulses of UV and visible radiation. In this technique, the absorbance change induced in the sample solution by the UV or visible exciting flash is monitored by an analyzing light beam passing through the sample and reaching a detector (photomultiplier or diode) at the selected wavelength via a monochromator. The detector senses changes in the analyzing light intensity and converts them into electrical signals in order to display the temporal history of the absorbance in a computer. Laser flash photolysis experiments can determine the ability of light of different wavelengths (UV, visible, and NIR) to initiate hybridization reactions of nucleic acid molecules in aqueous solutions and gels and on solid surfaces, the most effective wavelength for initiating hybridization reactions, the effect of laser light on reaction rates, the mechanisms and kinetics of the hybridization process and the effects of temperature, concentration, pH, and additives, e.g., buffer and enzymes, on reaction rates. Results from such analyses can be used to control the reaction mechanism. For example, NIR flash photolysis may be used to measure the hybridization kinetics that cover the wavelengths from about 850 nm to about 970 nm.

RNSDOCID: <WO 0047600A1 IA>

Development of Chip Technology - Surface Chemistry of Nucleic Acid Molecules on Semiconductors

The experimental verification for demonstrating the self-assembly of nucleic acid molecules on a solid semiconductor substrate, such as silicon and GaAs, may be conducted as described by Ratner *et al.* (Boland, T. and Ratner, B.D., *Langmuir 10*:3845-3852 (1994); Park, D., *et al.*, *Macromolecules 28*: 2595-2601 (1995)). The morphological coplanar patterns of the self-assembled monolayer nucleic acid molecules may be achieved by either one or both of the following techniques: (1) using a deep UV photolithographic method to produce high-resolution spatial regions and (2) introducing an anchoring functional group onto the nucleic acid molecules to enhance the spontaneous self-assembly in a two-dimensional array on a solid substrate.

The use of a deep UV photolithographic method to produce high-resolution spatial regions depends on the photochemistry of the nucleic acid molecules and substrates. The mechanism of the surface photochemistry involves the identification of the nucleic acid molecule free radicals and their reaction kinetics. High-resolution spatial regions having different surface free energies for controlled adhesion of various nucleic acid molecules can be created using deep UV photolithographic techniques. It has been demonstrated that, upon deep UV irradiation, a properly chosen self-assembled monolayer (SAM) undergoes a photolytic cleavage reaction (Boland, T. and Ratner, B.D., *Langmuir 10*:3845-3852 (1994)). This principle is used to create spatially different areas of reactivity that may be amenable to re-modification with another SAM forming moiety. The ability to modify spatially the self-assembled monolayers may be important for high-density memory.

In the introduction of an anchoring functional group to nucleic acid molecules to enhance spontaneous self-assembly in a two-dimensional array on solid substrates, various anchoring groups may be introduced to the nucleic acid molecules to achieve better adhesion to the substrate. An anchoring functional group such as thiol or silane may be introduced to nucleic acid molecules to

enhance the spontaneous self-assembly in a two-dimensional array on solid GaAs, for example. These functional groups strongly bind the nucleic acid molecules to the substrate surface thereby immobilizing it. The two-dimensional order is enhanced whenever the regular structure of the adsorbed molecule corresponds to the structure of the underlying substrate.

The mechanism of adsorption of DNA bases on to a gold surface has been studied by Boland and Ratner (Boland, T. and Ratner, B.D., *Langmuir 10*:3845-3852 (1994)) who have shown that the DNA bases self-assemble on crystalline gold in an ordered two-dimensional lattice. The process of self-assembling on gold was characterized by adsorption and surface migration which progressed in three stages. The initial appearance of random clusters of ordered molecules at the gold reconstriction sites was observed, i.e., some of the adsorbed molecules may stick to favorable, active sites on the gold while other diffusing molecules cluster with the ones initially adsorbed to build islands until the active site is completely occupied. The islands are clusters of molecules held in place by intermolecular force and adsorbed onto an active site on the gold. Only a limited number of molecules are allowed to be adsorbed onto active sites, occupation of all reconstruction sites by the clusters and rearrangement of the molecules to cover the remaining gold surface with a two-dimensional lattice. After step two is completed, the clusters have set up an epitaxial structure from which a wave of crystallization into a two-dimensional lattice can occur. The crystallization wave progresses perpendicular to the chains. When the substrate surface is GaAs, both the S-H group in thiol and the N-H group in the DNA bases may be functional groups. Scanning tunneling microscopy (STM) and electron spectroscopy for chemical analysis (ESCA) may be used to investigate the adsorption of nucleic acid molecules on GaAs.

RNSDOCID: <WO 0047600A1 IA>

Development of Chip Technology - Algorithm Development and Application

The use of nucleic acid in memory storage systems requires algorithm development and error analysis in order to reliably store and retrieve the data. Algorithm development and error analysis entails developing models of data storage using the unique properties of nucleic acid such as DNA. Exemplary memory models include simple ones derived by analogy with electronic memory devices to more complex models capable of massively parallel database searches. Models of other classical computing problems such as data sorting may also be employed. Preferably, the memory storage devices are based on chips containing immobilized nucleic acid molecules in a gel of mobile target nucleic acid molecules. The chips are etched and the immobilized nucleic acid molecules are attached in addressable arrays. Addressable array chips are commonly used for DNA Sequencing by Hybridization (SBH) (Pease, A.C., *et al., Proc. Natl. Acad. Sci. USA 91*:5022-5026 (1994); Lipshutz, R., *et al., BioTechniques. 19*:442-447 (1995); Noble, D., *Anal. Chem. 67*:201A-204A (1995); Frutos, A.G., *et al., Nucl. Acids Res. 25:* 4748-4757 (1997)).

The concept for a nucleic acid memory chip is straightforward. There are two operations to be performed - storage and retrieval. There are specific difficulties associated with the storage and retrieval of addressable data in aqueous DNA solutions (Baum, E.B., *Science 268*: 583-585 (1995)), therefore, an addressable surface array is a desirable starting point for developing a more general three-dimensional storage system. Starting with surface-attached nucleic acid molecules, it may be possible take advantage of exotic topological structures that can be constructed with DNA (Chen, J. and Seeman, N.C., *Nature 350*:631-633 (1991); Seeman, N.C., *Accounts of Chem. Res. 30*:357-363 (1997); Winfree, E., *DIMACS Series in Discrete Mathematics and Theoretical Computer Science 27*:199-221 (1996)) in the form of three-dimensional memory storage systems.

The first model of nucleic acid memory to be considered is bit, byte, or word storage in direct analogy with digital electronic memory. When each pixel

SDOCID: <WO___0047600A1_IA>

on the chip represents a single bit, then a chip with 4^{15} pixels, for example, could store 1 Gigabit or 128 Megabytes. This could be quite cumbersome, however, inasmuch as each pixel in a memory area would have to be queried to read the data. If the data is stored as bytes, rather than bits, then 2^8 (256) pixels are required to represent one byte. This reduces the number of pixels to be interrogated, but it reduces the chip capacity correspondingly, to a mere four Megabytes. Likewise, going to larger memory units, e.g., words, again reduces the available memory. However, word strategies minimize the noise level in the system and maximize the stability of hybridized complementary pairs.

In an alternative model, each pixel need only represent a 0 or 1 state. Therefore, fabrication of the chip is greatly simplified. In fact, all the pixels could contain the same nucleic acid molecules. Likewise, the composition of the target nucleic acid molecules is very simple. The disadvantages of this system are the tradeoffs in storage capacity versus the number of storage and retrieval operations needed, e.g., the number of pixels to be hybridized or fluoresced.

Another alternative model of nucleic acid memory is derived from DNA Sequencing by Hybridization (SBH). In this method, a nucleic acid fragment to be analyzed is incubated with a large set of oligonucleotide probes of known sequence. The extent to which the analyte nucleic acid molecule hybridizes with each probe is measured to identify those probes that perfectly complement the nucleic acid subsequences. This information can then be used to identify the analyte sequence. For example, this may be achieved by using detectably labeled, e.g., fluorescently-tagged, target nucleic acid molecules and reading the addresses of the labeled, e.g., fluorescing, pixels when the chip is illuminated. Recent experiments have illustrated hybridization to probes synthesized in 8-10 μm site with 4^{11} oligonucleotides placed in an array measuring 2-cm square. Array feature resolutions of about 1-2 μm are projected with advances in UV photolithography. Because of the combinatorial strategies used in fabricating the chip, the set of all 4^k oligonucleotides (k-mers) can be sequenced in 4 x k synthesis cycles (Pease, A.C., *et al.*, *Proc. Natl. Acad. Sci. USA 91*:5022-5026 (1994)). The number of

3NSDOCID: <WO___0047600A1_IA>

unique probes increases exponentially with *k* while the number of steps increases only linearly

The same concept can be applied to memory storage by encoding a data string in a base-4 code over the DNA alphabet set (A, C, G, and T). The data string is then a virtual DNA strand that can be stored on the chip by activating the pixels corresponding to its complementary subsequence. This assumes that the target medium has a complete set of fluorescence-tagged complementary subsequences to the probe strands. As an example, three bytes are stored in an array consisting of octanucleotide (8-mer) probes. *See* Fig. 5. Each byte requires a string of four characters from the base-4 alphabet set (A, C, G, and T). So three bytes translate to a 12-mer hypothetical DNA strand.

Ideally, only perfectly matched complementary polynucleotides hybridize. However, partial and slide matches are not only possible, but often cannot be avoided Moreover, the stability of the hybridized nucleic acid molecules is dependent upon the percentage C/G content in the nucleic acid molecules so that even the perfect matches will vary in concentration (Smith, L.M., *et al.*, *J. Comput. Biol. 5:* 255-267 (1998)) More complex DNA word strategies have been suggested, (Frutos, A G , *et al.*, *Nucl. Acids Res. 25:* 4748-4757 (1997)), but these lead to longer DNA strands and other complications associated with longer strands. Others have also pointed out that long strands are required to assure that all the nucleic acid molecules and their complements are distinct (Frutos, A.G., *et al.*, *Nucl. Acids Res. 25:* 4748-4757 (1997)). However, experimental results for SBH have shown that fluorescence signals from complementary nucleic acid molecules are 5-35 times stronger than those with single or double base-pair mismatches which demonstrates specificity in the identification of complementary sequences (Pease, A.C., *et al.*, *Proc. Natl. Acad. Sci. USA 91:*5022-5026 (1994)).

An alternative model of nucleic acid memory is one that exploits the massive parallelism that is well-known in other areas of DNA computing, particularly in the areas of combinatorial problems and data encryption. In this model, the chip is not used for writing to memory, but rather contains a permanent

database or library in which each pixel is encoded with different data. The database can be "searched" by populating the target medium with synthesized nucleic acid strands, preferably DNA strands, encoded with the search criteria. Only those sites containing the criteria will hybridize. The hybridization reactions may be initiated either by flashing the entire chip with laser light or by mixing an enzyme with the nucleic acid target strands. By illuminating the chip and using the fluorescence methods used in SBH, the results can be retrieved. For example, in searching for a fingerprint match, by encoding an unknown fingerprint into the base-4 alphabet set (A, C, G, and T) and synthesizing and replicating a DNA molecule to place in the target medium, the entire database can effectively be searched simultaneously (massive parallelism). Likewise, by using different fluorescence types, the search can be carried out for several pieces of data concurrently.

The following examples are illustrative, but not limiting, of the methods of the present invention. Other suitable modifications and adaptations of the variety of conditions normally encountered which are obvious to those skilled in the art are within the spirit and scope of the present invention.

Examples

Example 1 - Hybridization of an oligonucleotide in aqueous solutions induced by UV, Visible, and NIR light

Experimental Protocol

Materials and Solutions: An oligonucleotide with 27 bases (5'- CGA CGA CGA CGA CGA CGA CGA CGA CGA -3') (SEQ ID NO:1) was purchased in highly purified form (analytically estimated by HPLC and capillary electrophoresis) from Oligos Etc., Inc. (Wilsonville, OR). This particular sequence was used because an individual molecule has a low natural tendency for self-hybridization as every third base (A) is mismatched with a potential complementary strand (stem loop melting temperature of 0°C) and it does not contain thymine. *See* Fig. 14.

BNSDOCID: <WO 0047600A1 IA>

Thymine was excluded to prevent the thymine-thymine cyclobutane dimers upon irradiation with UV light.

Aqueous solutions containing 3.7 μM single stranded oligonucleotide (27 base pairs) were prepared at three pH values in different buffers: pH 4.5 (10 mM acetic acid), pH 7.8 (10 mM potassium phosphate), and pH 10.1 (10 mM sodium carbonate/sodium bicarbonate). All solutions were autoclaved at 121°C for 30 minutes prior to the addition of the oligonucleotides.

Irradiation: Photolysis was performed at several wavelengths. Irradiation at 450 nm and 920 nm was performed with an argon laser (Coherent Laser group, Santa Clara, CA) with an intensity of 4.45 mW/cm^2 at both wavelengths at exposure times of 0.2, 2.0, and 15 seconds. Irradiation at 300 nm was performed with an Hg lamp with an intensity of 14 mW/cm^2. Exposure times were 0.1, 0.6, and 4.0 seconds. The intensity of the visible source (430 nm) was 1.0 mW/cm^2, and exposure times were 0.4, 4.0, and 30 seconds. Laser irradiations were carried out at a pulse-width of 100 femtoseconds with an intensity of 4.0 mW/cm^2 at three wavelengths (450, 514, and 621 nm).

Analytical: The photolyzed samples were analyzed using precast polyacrylamide gels (15% polyacrylamide in TBE (tris-borate-EDTA)) from BioRad Co., (Hercules, CA). Samples (8.3 μL) were combined with 2 μL of gel loading solution (Sigma, type 1, gel loading buffer, 6X concentrate) prior to loading. The DNA ladder (Life Technologies, Inc.) (1.0 *ug/ul)* consisted of thirty-three 10 bp repeats and was suitable for sizing both single-stranded and double-stranded DNA fragments from 10 bp to 200 bp. The electrophoresis measurements were made with the ladder in the gel.

The gels were run in 1X TBE buffer (0.1 M Tris, 0.09 M boric acid and 1 mM EDTA, pH = 8) at low temperature (4°C) and low voltage (5 V/cm) to prevent denaturing of small DNA fragments by heat generated by the passage of electric current. DNA was visualized by ethidium bromide staining (UV transilluminator). Images were captured and stored via CCD camera (Eagle-Eye, Strategene). Band quantities were evaluated using NIH Image (v1.6). Error

analysis was performed on identical samples and on repeated experiments and the standard error was reported.

Results

Gel electrophoresis results for solutions irradiated at 300 nm (A), 450 nm (B), and 920 nm (C) are shown in Figs. 6A-6C. Gel electrophoresis images of the oligonucleotides before and after photo-induced hybridization in neutral aqueous solutions are shown. A strong band of the original 27 base oligonucleotide single-stranded is detected along with a weak slower-migrating band of a 54-base oligonucleotide. *See* Figs. 7A-7C. The quantity of dimer, as indicated by the intensity in the ethidium bromide stained polyacrylamide gels, is much higher in the photolyzed samples. For example, in samples photolyzed for 0.1 seconds at 300 nm, 27% of the original single-stranded oligonucleotide was converted to the higher molecular weight form. Similar results were obtained for other samples photolyzed for 15 seconds at 450 nm and 920 nm.

To examine whether the dimer was formed by hydrogen bond hybridization or by covalent bonding, the samples (9.0 μl) were heated to 95°C for 10 minutes to dissociate hydrogen bonds. After heating, only a trace of the dimer remained. Based on relative band intensities, the sample photolyzed for 0.1 seconds at 300 nm (Fig. 6A) showed that 26% of the original single-stranded oligonucleotides had been converted into the higher molecular weight form, but after heating, only a trace of dimer remained (approximately 3%). *See also* Fig. 7A. Similar results were obtained for other samples photolyzed at 450 and 920 nm (Figs. 6B and 6C). *See also* Figs. 7B and 7C. Additionally, 8 μl samples were treated with 7 μl NaOH solution (pH 13) which resulted in complete dissolution of the higher molecular weight band.

To examine whether the intensities of the bands in Figs. 7A-7C increase linearly with the weight of the oligonucleotide, various concentrations of the oligonucleotide were run in the same gel. For each concentration a control

ENSDOCID: <WO 0047600A1 IA>

(unexposed) sample was also processed on the gel. The sum of the intensities of the two bands, the 27- and 54-mer bands, was plotted against the weight of the oligonucleotide. A typical plot is shown in Fig. 8. In all cases, the band intensity increased linearly with oligonucleotide weight up to an intensity of 8×10^3. Therefore, any band intensity higher than this value was discarded.

In addition, DNA melting experiments were performed. The absorbance at 260 nm of 100 μl of photolyzed samples (pooled from several experiments) was measured between 20°C and 90°C in 5-20°C increments (Figs. 9A and 9B). Aliquots were removed at each temperature and analyzed by polyacrylamide gel electrophoresis. As the temperature increased, the upper band intensity decreased and the solution A_{260} increased monotonically, which is typical for DNA hybrid melting (A_{260} ss≈ 1.5 A_{260} ds). The theoretical melting temperature for a perfect match 27mer based on similar %GC content is 83.8°C (Oligos Etc. Inc., Wilsonville, OR) As noted in the Experimental protocol section, the oligonucleotide used here was constructed for third base pair mismatch when two single-stranded oligonucleotides hybridize, hence the theoretical melting temperature for the 27mer employed is 40-50°C. The observed melting temperature (≈60°C) is bounded by these extremes. These results constitute firm evidence that photo-induced hydrogen bond dimerization was achieved rather than covalent bonding.

Table 1 shows a calculated "efficiency" of the UV (300 nm) photo-induced hybridization at pH ≈7.8 and at various oligonucleotide concentrations. This efficiency is defined as the intensity of the upper band (dimer) divided by the sum of the upper and lower band intensities. This result was only recorded for samples shown to be in the linear range of the calibration solutions and the gels (quantity < 160 ng per band). Namely, the intensity of the ethidium bromide signal of a single band was linear with 4.0g DNA up to 8×10^3 intensity units. Data from lanes with intensities above this limit were not included. At 300 nm, an overall efficiency of 27 ± 3.2% was found after photo-induced hybridization, compared with 10.6 ± 2.5% in the unirradiated controls.

ISDOCID: <WO___0047600A1_IA>

Table 1

Efficiency of Photo-Induced Hybridization of Oligonucleotides at Various DNA Concentrations

Sample pH = 7.8	Amount of DNA (ng)	Upper Signal	Lower Signal	Total Signal	Efficiency (%)	
Photolyzed (0 1 sec @ 300 nm)						
1	250	2306	5987	8293	27.8	
2	250	2542	7143	9685	26.3	
3	129.6	1441	5123	6564	22.0	
4	97.5	1163	2953	4116	28.3	Avg.: 27.0±3.2
5	54.9	841	1913	2754	30.5	
Unirradiated sample[a]						
1	250	898	6742	7640	11.8	
2	250	622	7155	7777	8.0	
3	250	1164	7810	8974	13.0	
4	250	965	7566	8531	11.3	
5	250	696	6431	7127	9.8	
6	129.6	687	6353	7040	9.8	
7	129.6	451	4712	5163	8.7	Avg.:
8	129.6	662	5904	6566	10.1	10.6±2.5
9	91.5	560	4271	4831	11.6	
10	91.5	385	4176	4561	8.4	
11	91.5	385	5045	5430	7.1	
12	54.9	411	2162	2573	16.0	
13	54.9	312	3091	3403	9.2	
14	54.9	576	3553	4129	14.0	

[a]data are obtained from three different gels.

Additional experiments were run at various pH and wavelengths and hybridization ratios were calculated. Specifically, each sample (control and irradiated) contained the same amount of DNA and the bulk of the DNA was in the single-stranded form. The intensity of the upper band (dimer) from irradiated samples was divided by the intensity of the upper band from controls (after

RNSDOCID: <WO 0047600A1 IA>

normalization by lower band intensity) because this ratio between double-stranded hybrids is a more sensitive indicator of the effect of photolysis. In Fig. 10, the effect of pH on the HR (300 nm, exposure time of 0.1 seconds) is shown. This was not surprising as pH extremes leave ionic interactions more unlikely due to the predominance of protonated or deprotonated forms. Also, in Fig. 10, the relative HRs for all samples at three wavelengths are shown. Although UV (300 nm) irradiation was most effective, significant hybridization was formed in the NIR range (920 nm). The values of the hybridization ratios were found to be 3.5 and 2.5 at 300 nm and 920 nm, respectively.

The effect of pH on the hybridization ratio at 300 nm with an exposure time of 0.1 seconds was investigated. The hybridization ratio decreased from about 3.0 at a pH of about 7.8 to about 1.0 at a pH of about 4.5 and about 0.6 at a pH of about 10.1. Fig. 11 shows the effect of the exposure time with 300 nm light on the hybridization yield of 37 μM at pH 7.8. Hybridization yield increased to a maximum value at 0.1 seconds. At higher exposure times (0.6 and 4.0 seconds), UV photolysis appears to induce scission of the hybridized molecules. Moreover, the relative hybridization ratios of the DNA sample (pH= 7.8) photolyzed with a broad spectrum band which includes UV, VIS and IR regions are shown in Fig. 12. In these experiments, both, steady state and flash photolysis techniques were utilized. As expected, UV irradiation (300 nm) induced more hybridization than the visible and IR irradiation.

Overall, photo-induced hybridization was successfully achieved. The results show that irradiation of a single-stranded DNA by light with wavelength of 300, 450, and 920 nm, induces hybridization between complementary bases. Covalent dimerization was prevented by the absence of a thymine moiety. The photolysis improved the hybridization reaction especially in the vicinity of pH of 7.8. That is, aqueous solutions of DNA with pH of about 7.8, irradiated by UV at 300 nm for 0.1 seconds, gave rise to the highest relative extent of hybridization.

Discussion

The results demonstrate considerable hybridization of the oligonucleotides as a result of UV irradiation. Moreover, visible and NIR irradiation led to hybridization, although to a lesser extent. Since hybridization involves the dissociation of hydrogen bonds (between a base and water molecule) and the formation of other hydrogen bonds (between complementary bases), all the wavelengths used in this study had sufficient energy to effect such changes.

Since UV light is absorbed by the DNA bases, an additional mechanism appears to be operating in this case as a result of direct excitation of the oligonucleotide The excitation is absorbed by the oligonucleotide and the excited state can undergo charge separation to produce an anion radical and a cation radical of different bases. It is known that radiolysis and photolysis of DNA produce cation radicals, mainly $A^{\bullet+}$ and $G^{\bullet+}$, and anion radicals, mainly $C^{\bullet-}$ and $T^{\bullet-}$, due to the relative electron affinities of the DNA bases. Steenken (Steenken, S., *Free Rad. Res. Commun. 16*:349-379 (1992); Steenken, S., *Biol Chem. 378*:1293-1297 (1997)) showed that the anion radicals react rapidly with protons and the cation radicals can readily donate protons. As a result, if $C^{\bullet-}$ of one oligonucleotide approaches $G^{\bullet+}$ of another nucleotide, rapid proton transfer between them will take place to produce two neutral radicals that are hydrogen-bonded. These radicals are eventually repaired, for example, by charge migration and recombination, but the repaired bases can remain hydrogen bonded. The findings that UV-induced hybridization is more effective at pH 7.8 than at pH 4.5 or pH 10.1 may be interpreted to support this mechanism, i.e., when the medium is more likely to supply the proton to $C^{\bullet-}$ (at low pH) or more likely to remove a proton from $G^{\bullet+}$ (at high pH), then the probability of proton transfer between these two species is diminished and thus the extent of hybridization is decreased.

With visible and NIR light, for which charge separation is unlikely, the principal path whereby the hybridization occurs remains unclear. As suggested above, a reasonable postulate is photon-induced breakage of hydrogen bonds

3NSDOCID: <WO__0047600A1_IA>

between the nucleotide and water molecule followed by the exothermal hydrogen bonding of complementary bases. Despite the higher efficiency of UV light in effecting hybridization, it is more practical to use visible or NIR light sources for analytical and/or computer applications because of the avoidance of covalent (irreversible) linkage. Also, the observation that substantial hybridization occurred upon 920 nm irradiation is noteworthy as GaAs emission is typically at 920 nm. Thus, a possible memory device could contain a gel of oligonucleotide strands attached to a semiconductor wafer by appropriate end groups as described in detail above. Light emission at selected sites would cause hybridization at these sites. The extent of hybridization may be increased above the levels observed in this example by selecting oligonucleotide sequence and length (match/mismatch, GC content) (Ikuta, S., *et al., Nucleic Acids Res. 15*:797-811 (1987)). Also, the oligonucleotide concentration could be much higher for gene array and computer applications and the increased proximity of complementary strands may enhance hybridization.

The energy of a DNA hydrogen bond is typically about 0.35 eV. There is sufficient energy to dissociate one such bond not only by an ultraviolet (UV) photon (typically about 4.0 eV), but even by a near-infrared (NIR) photon (from about 1.0 to about 1.3 eV). Thus, even GaAs emission (about 1.3 eV) can, in principle, dissociate an adjacent pair of hydrogen bonds, thereby raising immediate prospects of erasing stored information.

Example 2 - NIR-Induced Hybridization

In these experiments, aqueous solutions of oligonucleotides (5' TGCA TGCA TGCA TGCA TGCA TGCA Y) (SEQ ID NO:2) were irradiated by NIR at 850 nm with gallium aluminum arsenide (GaAlAs) (Spectra Diode Laboratories) with the intensity of 76.4 W/cm^2. The sequence of this oligomer is self-complementary with no mismatches (dimer melting temperature 77.9°C). It also forms a stem loop structure (melting temperature 120°C based on the GC content).

Preliminary results demonstrate that irradiation with an 850 nm laser results in two distinct bands on polyacrylamide gels. The intensity of these bands varies with the type of buffer used for diluting the oligonucleotide. The following standard techniques were used to demonstrate that the upper band is indeed a double-stranded oligomer composed of two single-stranded hydrogen bonded oligonucleotides: (1) boiling the sample (5 minutes at 95°C) caused the upper band to disappear; (2) increasing the pH with NaOH (pH 13) also caused the upper band to disappear; (3) adding S I nuclease (which attacks single-stranded DNA specifically) significantly reduced the brightness of the lower band; and (4) measuring the UV absorption (260 nm) of the oligomer showed a 23% increase in absorption as the temperature was raised from 40°C to 90°C. The DNA melting curve determines the transition from double-stranded DNA into single-strand as shown in Fig. 13.

This invention has been described in specific detail with regard to specific materials and methods for the photo-induction of nucleic acid hybridization and applications thereof. Except where necessary for operability, no limitation to these specific materials is intended nor should such a limitation be imposed on the claims appended hereto. From the foregoing description, one skilled in the art can easily ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions without undue experimentation. All patents, patent applications and publications cited herein are incorporated by reference in their entirety.

BNSDOCID: <WO 0047600A1 IA>

What Is Claimed Is:

1. A nucleic acid memory device comprising a photo-excitation source, a substrate with nucleic acid molecules immobilized thereto and attached and in a combinatorial array, a mobile phase containing nucleic acid molecules at least partially complementary to the immobilized nucleic acid molecules, and a detector which detects hybridization within the combinatorial array.

2. The nucleic acid memory device of claim 1, wherein the photo-excitation source and the substrate are the same.

3. The nucleic acid memory device of claim 1, wherein the photo-excitation source is a semiconductor.

4. The nucleic acid memory device of claim 1, wherein the substrate is a semiconductor.

5. The nucleic acid memory device of claim 1, wherein the immobilized nucleic acid molecules are DNA molecules and the target nucleic acid molecules are DNA molecules.

6. The nucleic acid memory device of claim 1, wherein the detector is selected from the group consisting of: a confocal detector, a charged coupled device, and a flourescent tag.

7. The nucleic acid memory device of claim 6, wherein the flourescent tag is selected from the group consisting of: DIG-dUTP, fluorescein-12-dUTP, and 6-FAM phosphoramidite.

NSDOCID: <WO___0047600A1_IA>

8. The nucleic acid memory device of claim 1, further comprising a computer.

9. A method of inducing nucleic acid hybridization comprising:

(a) attaching a monolayer of single-stranded nucleic acid molecules in a combinatorial array of pixels to a substrate, wherein the nucleic acid molecules are immobilized;

(b) contacting the substrate with a mobile phase containing one or more target single-stranded nucleic acid molecules at least partially complementary to the immobilized nucleic acid molecules; and

(c) applying a photo-excitation source onto each pixel of the combinatorial array to induce hybridization between the immobilized nucleic acid molecules and the one or more target nucleic acid molecules.

10. The method of claim 9, wherein the photo-excitation source is a semiconductor.

11. The method of claim 10, wherein the photo-excitation source is a VCSEL array.

12. A method of inducing hybridization of two or more nucleic acid molecules which are at least partially complementary, comprising applying a photo-excitation source to the two or more nucleic acid molecules, whereby hybridization is induced.

13. A method for amplifying at least one double-stranded nucleic acid molecule, comprising:

(a) providing a first and second primer, wherein the first primer is complementary to a sequence at or near the 3'-terminus of the first strand of the nucleic acid molecule to be amplified and the second primer is complementary to a

INSDOCID: <WO___0047600A1_IA>

sequence at or near the 3'-terminus of the second strand of the nucleic acid molecule to be amplified;

(b) mixing the first and second primers, and the nucleic acid molecule to be amplified, with one or more polypeptides having nucleic acid polymerase activity thereby forming a mixture;

(c) illuminating the mixture with light under conditions favoring the photo-induced hybridization of the first primer to the first strand of the nucleic acid molecule, and the second primer to the second strand of the nucleic acid molecule;

(d) producing a third nucleic acid molecule complementary to the first strand and a fourth nucleic acid molecule complementary to the second strand;

(e) denaturing the first and third strands, and the second and fourth strands; and

(f) repeating steps (a) to (e) one or more times.

14. A method of sequencing a nucleic acid molecule, comprising:

(a) mixing a primer with a first nucleic acid molecule to form a first mixture,

(b) illuminating the first mixture with a photo-excitation source under conditions favoring the photo-induced hybridization of the primer with the first nucleic acid molecule;

(c) mixing the first mixture with one or more deoxyribonucleoside triphosphates (dNTPs), one or more polypeptides having nucleic acid polymerase activity and one or more terminator nucleotides thereby forming a second mixture;

(d) incubating the second mixture under conditions sufficient to synthesize a random population of DNA molecules complementary to the first nucleic acid molecule, wherein the synthesized nucleic acid molecules are shorter in length than the first nucleic acid molecule and wherein the synthesized nucleic acid molecules comprise a terminator nucleotide at their 3' termini; and

(e) separating the synthesized nucleic acid molecules by size so that at least a part of the nucleotide sequence of said first nucleic acid molecule can be determined.

15. The method of claim 14, wherein the one or more terminator nucleotides is selected from the group consisting of: ddTTP, ddATP, ddGTP, ddITP, ddCTP and mixtures thereof.

16. A method of inducing nucleic acid hybridization comprising:

(a) obtaining a substrate comprising a monolayer of single-stranded nucleic acid molecules in a combinatorial array of pixels immobilized thereon and a mobile phase containing one or more target single-stranded nucleic acid molecules at least partially complementary to the immobilized nucleic acid molecules; and

(b) applying a photo-excitation source to at least one pixel of the combinatorial array to induce hybridization between the immobilized nucleic acid molecules and the one or more target single-stranded nucleic acid molecules.

BNSDOCID: <WO__0047600A1_IA>

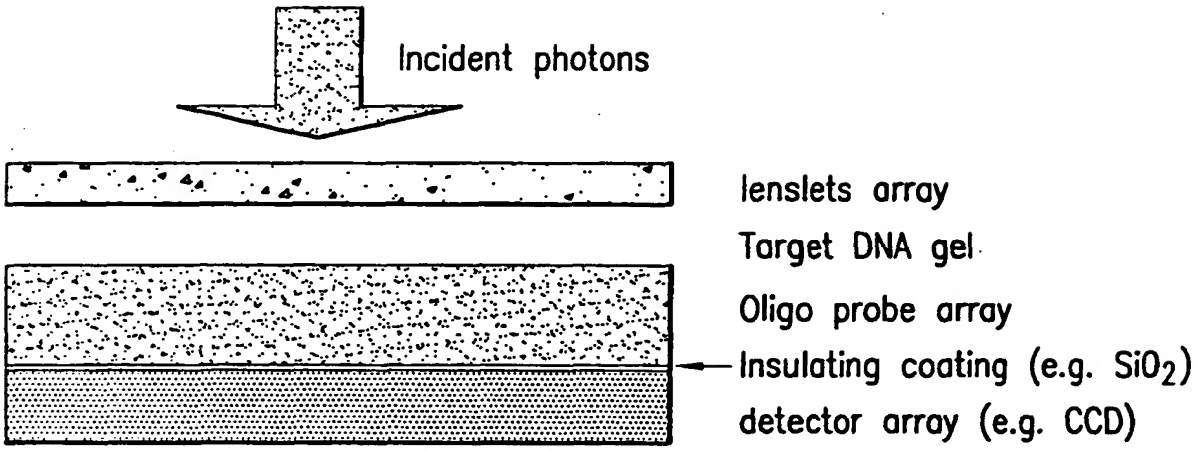

Incident photons
lenslets array
Target DNA gel
Oligo probe array
Insulating coating (e.g. SiO_2)
detector array (e.g. CCD)

FIG.1

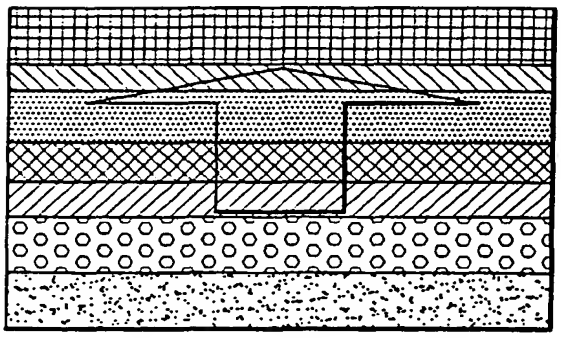

Detector array (e.g. CCD array)
Beam condensor (Lenslet array)
Target DNA gel
Oligo probe array
Beam conditioner (e.g. Lenslet array)
VCSEL array (e.g. GaN VCSEL array)
Substrate

FIG.2

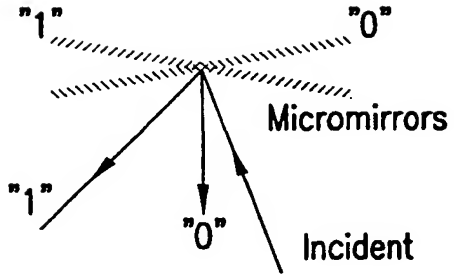


FIG.3A

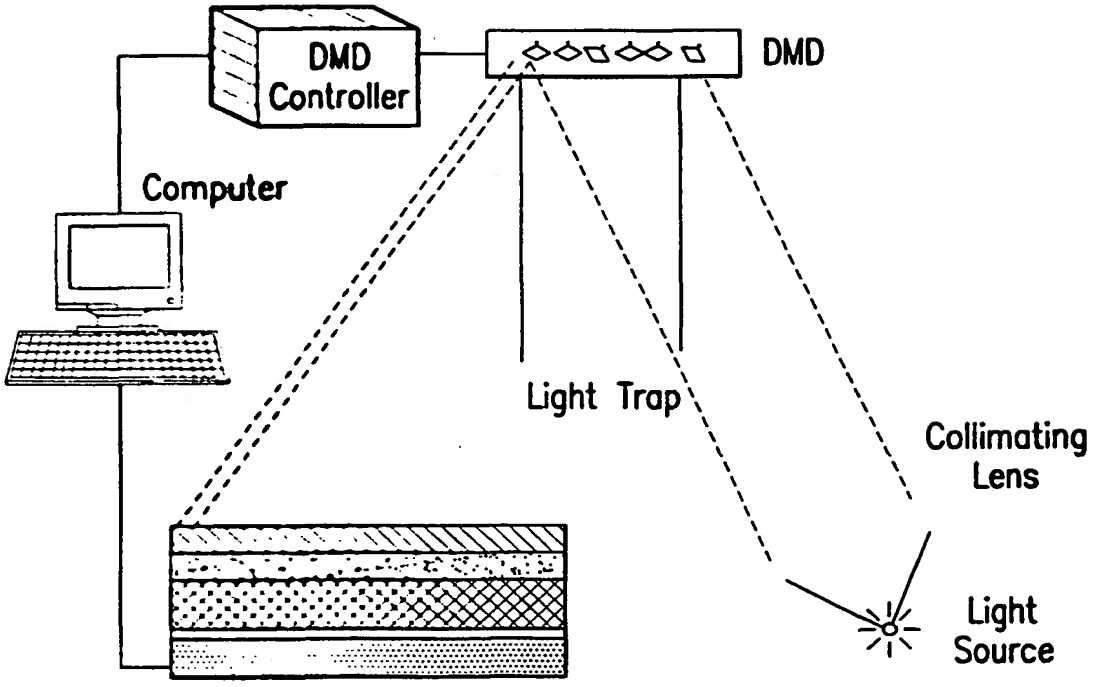


FIG.3B

3NSDOCID: <WO___0047600A1_IA>

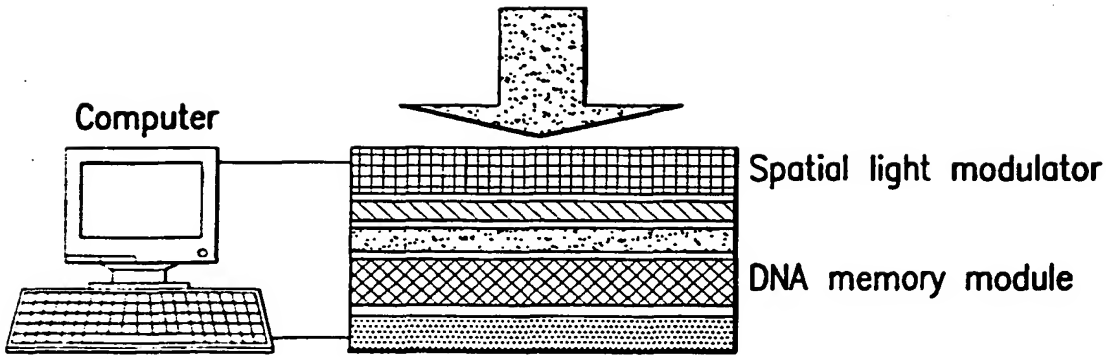


FIG.4

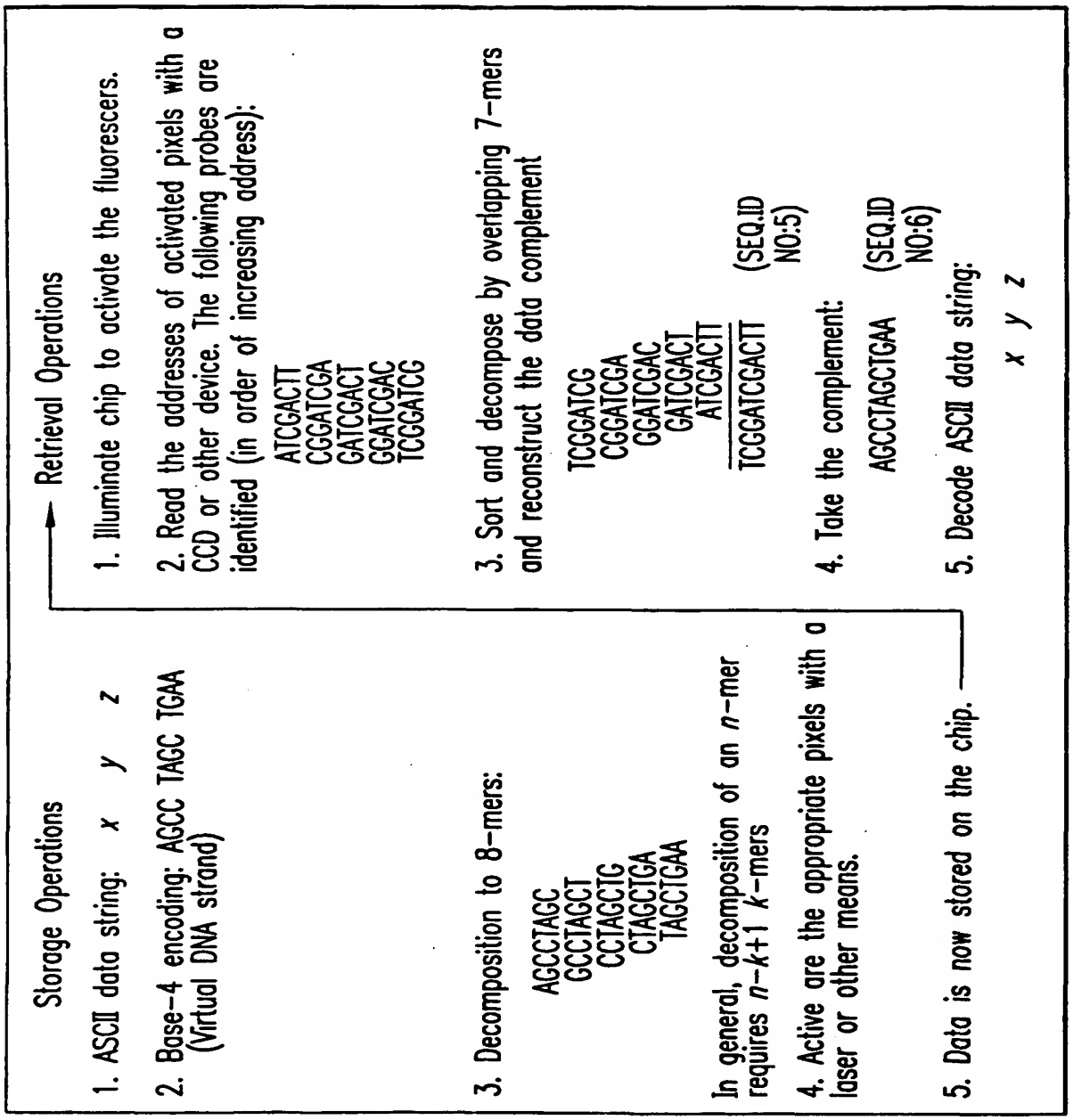

Storage Operations
1. ASCII data string: x y z
2. Base-4 encoding: AGCC TAGC TGAA (Virtual DNA strand)
3. Decomposition to 8-mers:
AGCCTAGC
GCCTAGCT
CCTAGCTG
CTAGCTGA
TAGCTGAA
In general, decomposition of an n-mer requires n-k+1 k-mers
4. Active are the appropriate pixels with a laser or other means.
5. Data is now stored on the chip.
Retrieval Operations
1. Illuminate chip to activate the fluorescers.
2. Read the addresses of activated pixels with a CCD or other device. The following probes are identified (in order of increasing address):
ATCGACTT
CGGATCGA
GATCGACT
GGATCGAC
TCGGATCG
3. Sort and decompose by overlapping 7-mers and reconstruct the data complement
TCGGATCG
CGGATCGA
GGATCGAC
GATCGACT
ATCGACTT
TCGGATCGACTT (SEQ.ID NO:5)
4. Take the complement:
AGCCTAGCTGAA (SEQ.ID NO:6)
5. Decode ASCII data string:
x y z

FIG.5

$\lambda_{300\ nm}$

Hybrid

Single strand

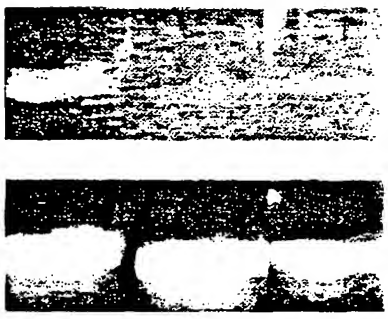


FIG.6A

$\lambda_{450\ nm}$

Hybrid

Single strand

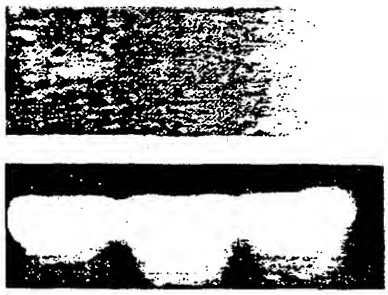


FIG.6B

$\lambda_{920\ nm}$

Hybrid

Single strand

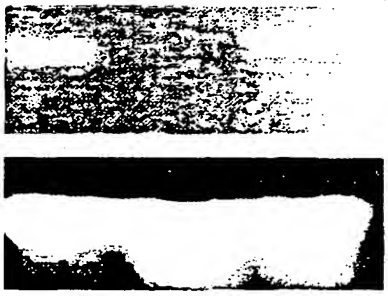


FIG.6C

Exposed Boiled Non-exposed

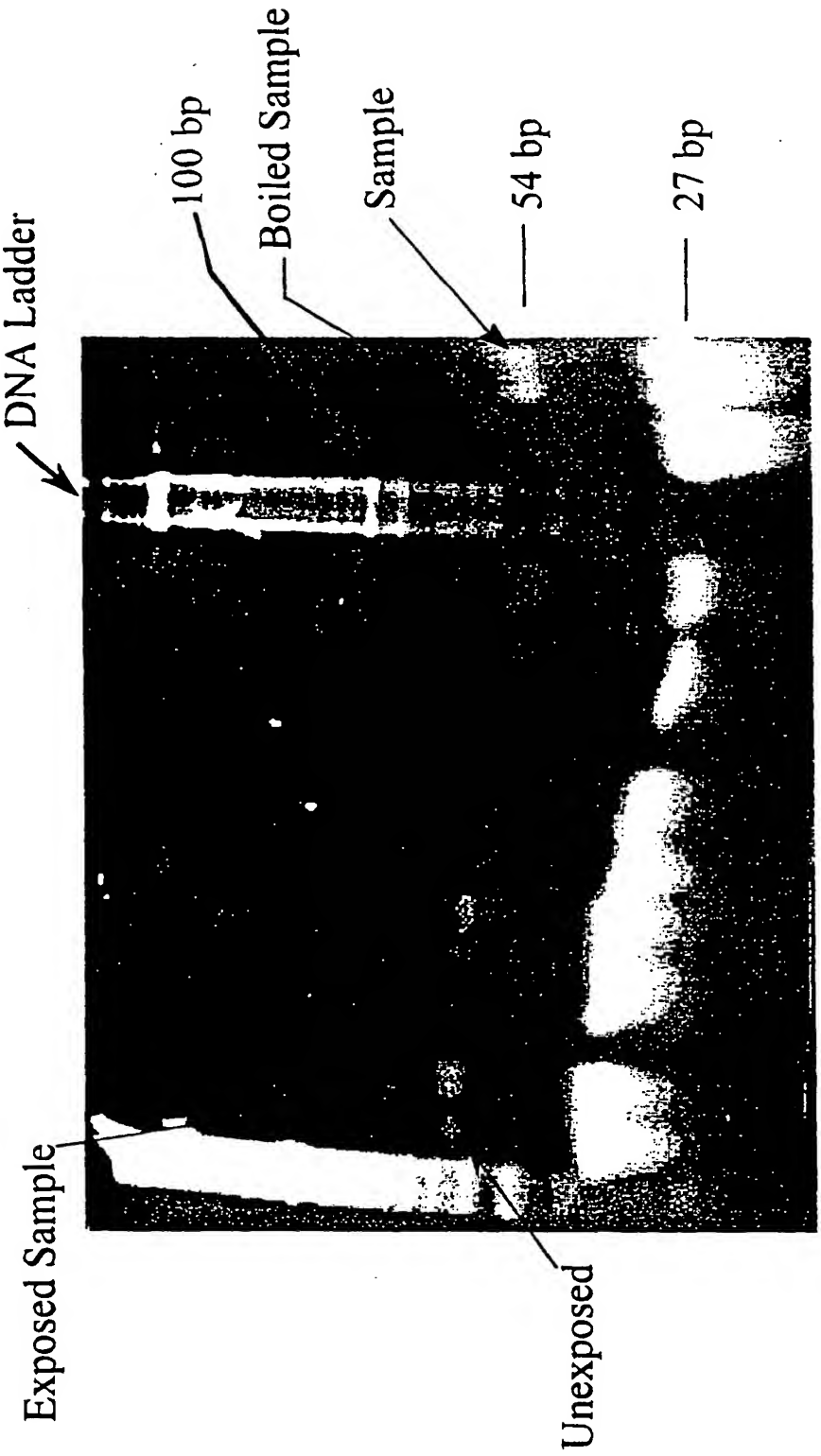

DNA Ladder
100 bp
Boiled Sample
Sample
54 bp
27 bp
Exposed Sample
Unexposed

FIG.7A

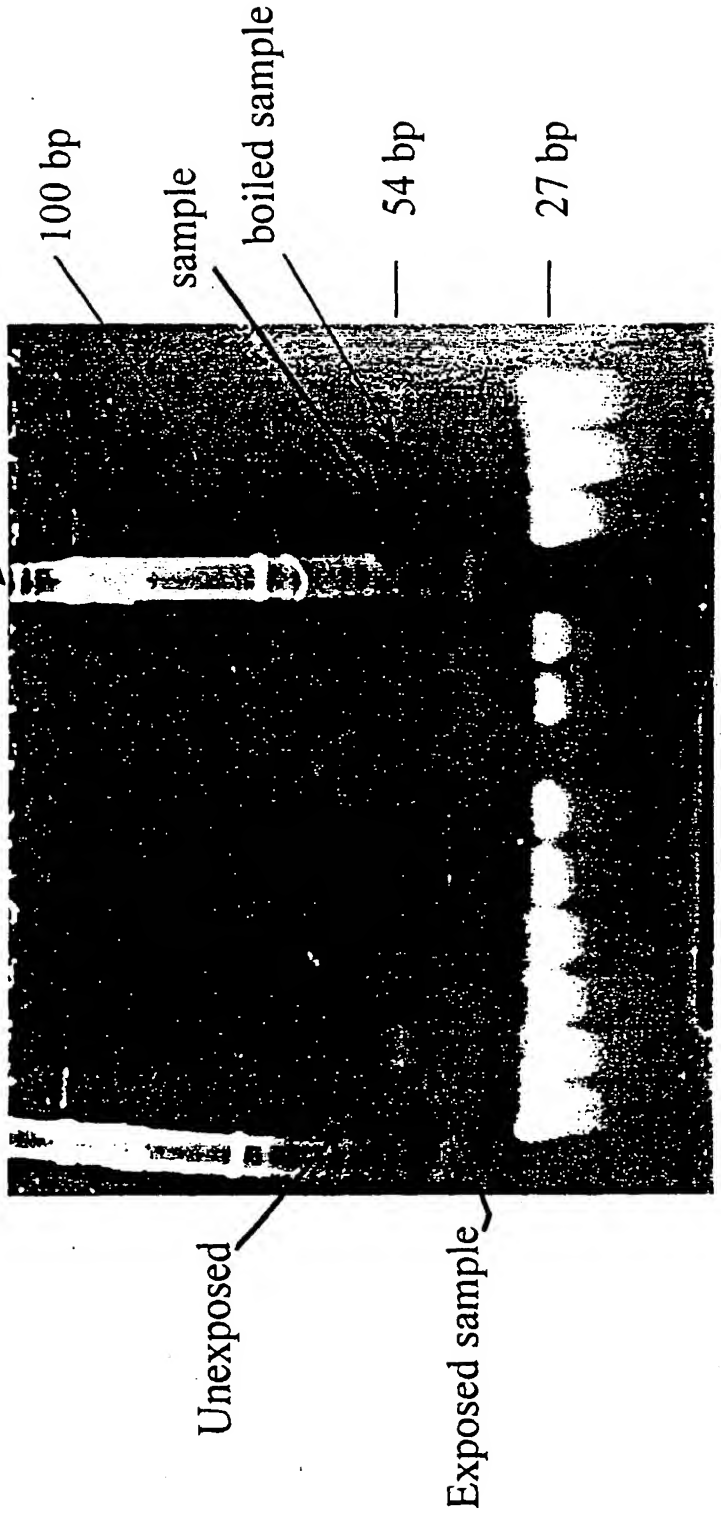

DNA Ladder
100 bp
sample
boiled sample
54 bp
27 bp
Unexposed
Exposed sample

FIG.7B

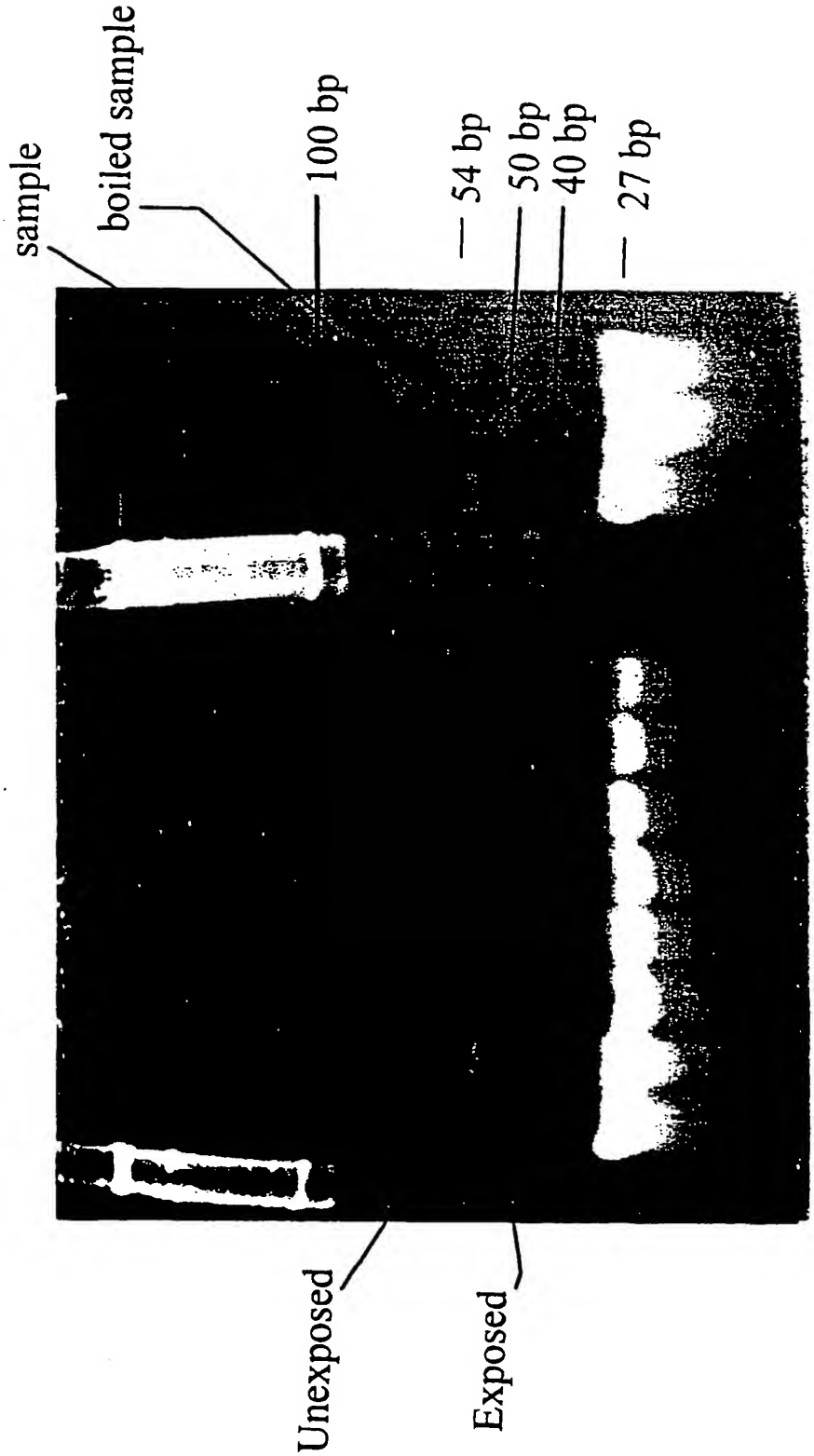

sample
boiled sample
100 bp
54 bp
50 bp
40 bp
27 bp
Unexposed
Exposed

FIG.7C

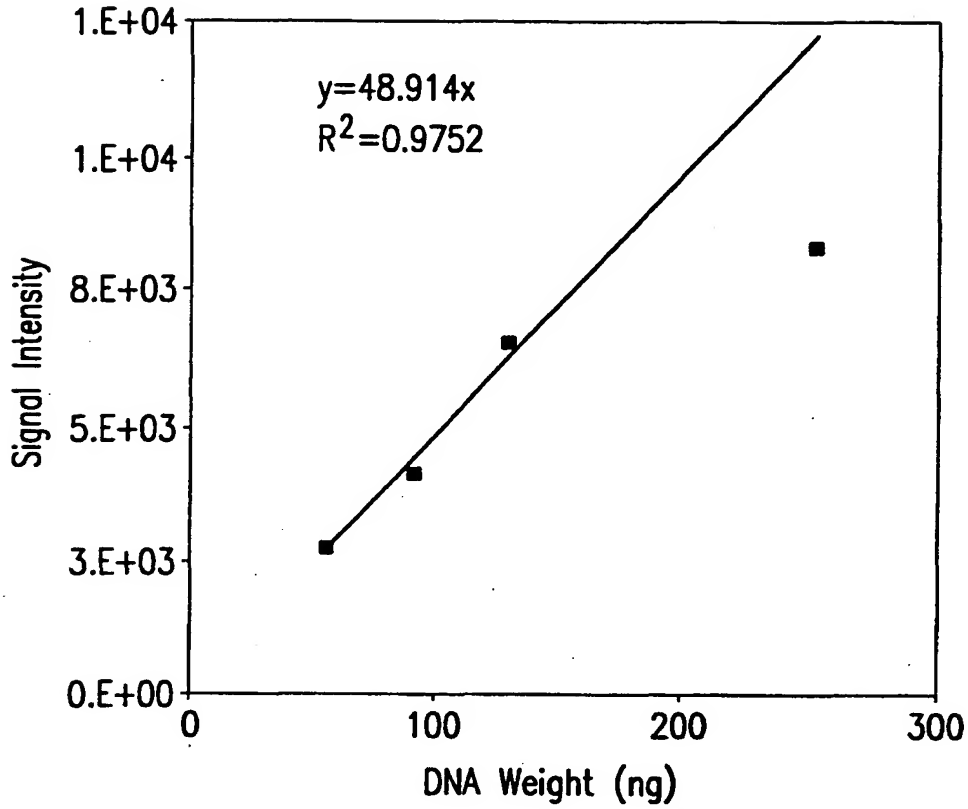

y=48.914x
R^2=0.9752
Signal Intensity
1.E+04
1.E+04
8.E+03
5.E+03
3.E+03
0.E+00
0
100
200
300
DNA Weight (ng)

FIG.8

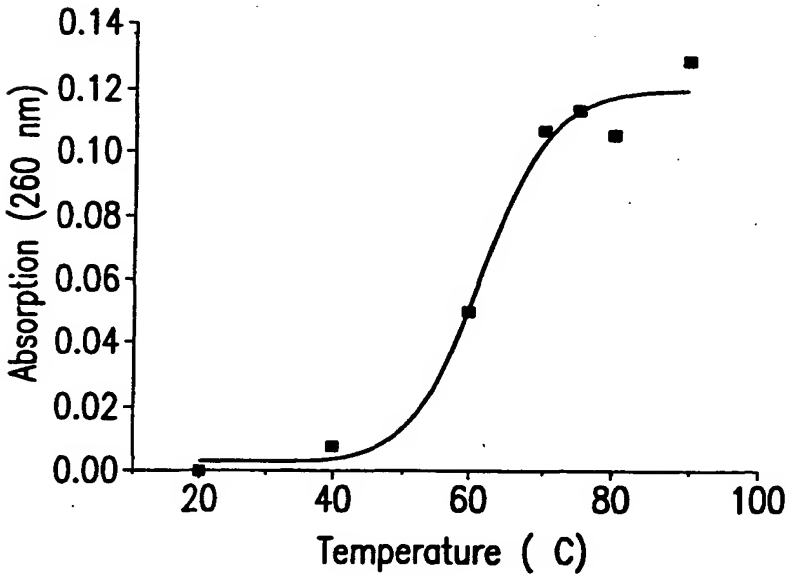

Absorption (260 nm)
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
20
40
60
80
100
Temperature (C)

FIG.9A

20 40 60 70 75 80 90

Temperature (C°)

FIG.9B

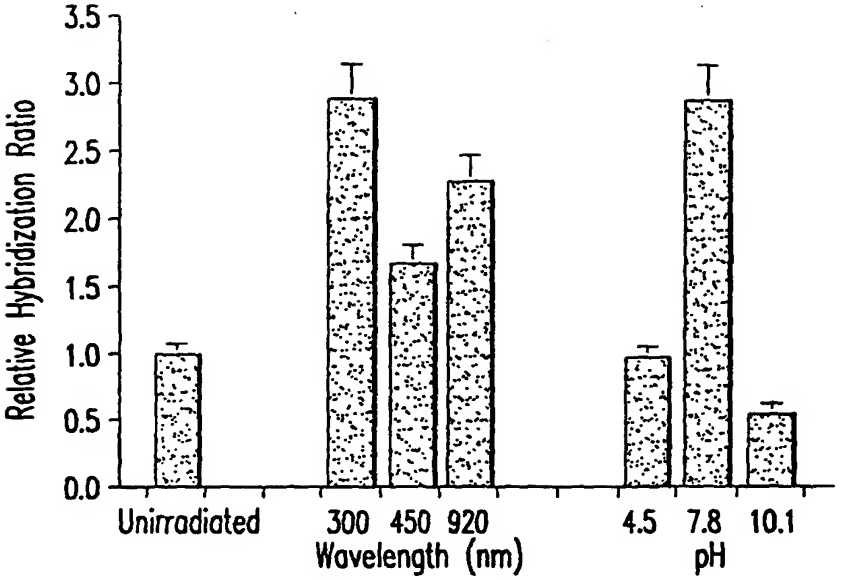

Relative Hybridization Ratio
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Unirradiated
300
450
920
Wavelength (nm)
4.5
7.8
10.1
pH

FIG.10

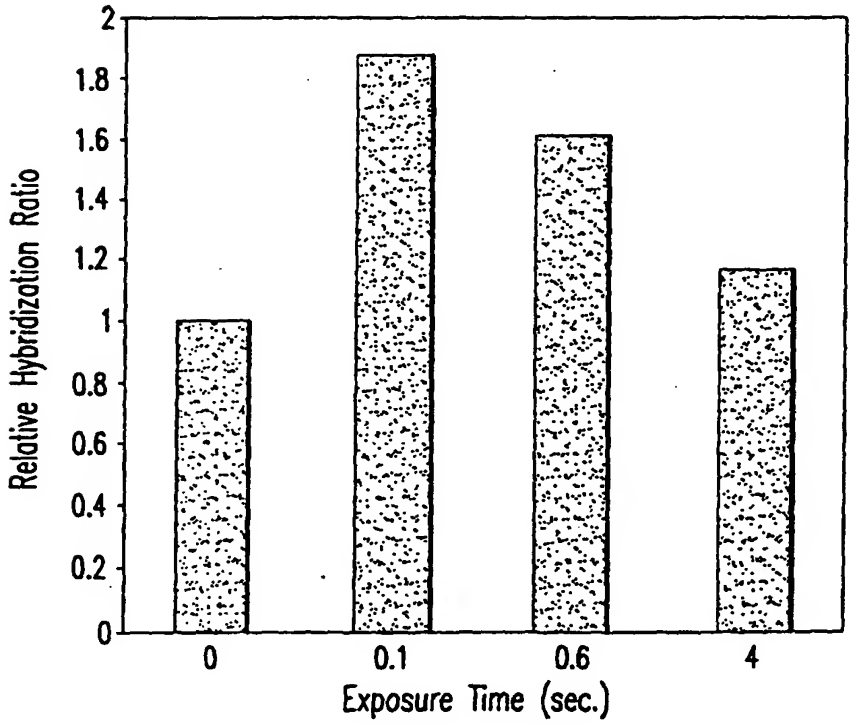

Relative Hybridization Ratio
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.1
0.6
4
Exposure Time (sec.)

FIG.11

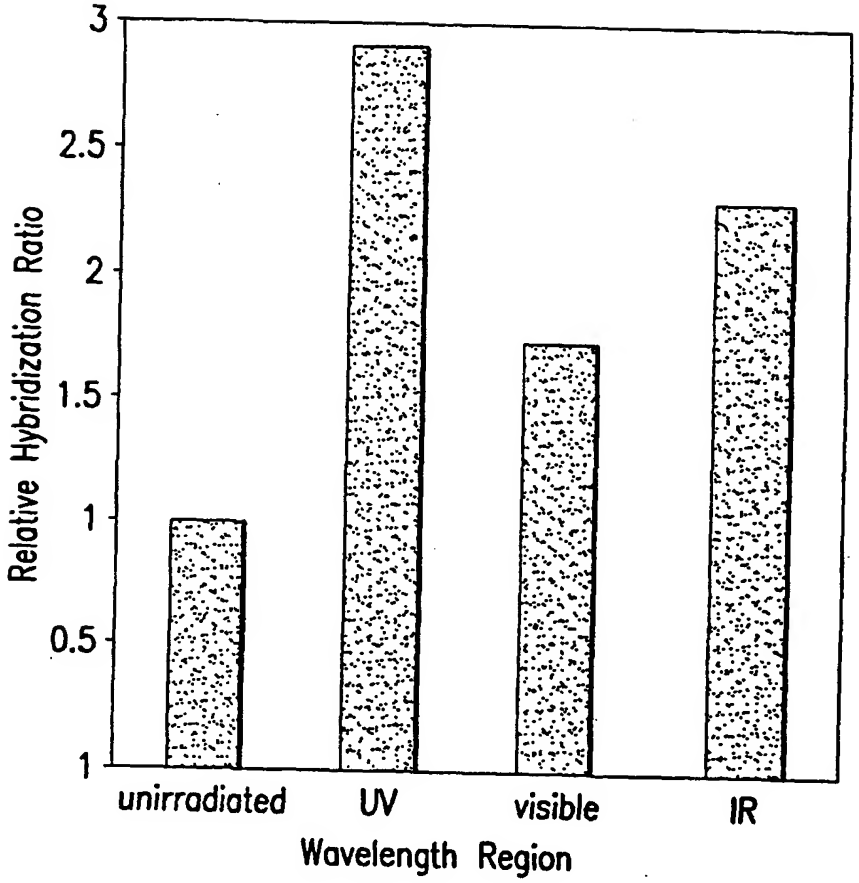

3
2.5
2
1.5
1
0.5
1
Relative Hybridization Ratio
unirradiated
UV
visible
IR
Wavelength Region

FIG.12

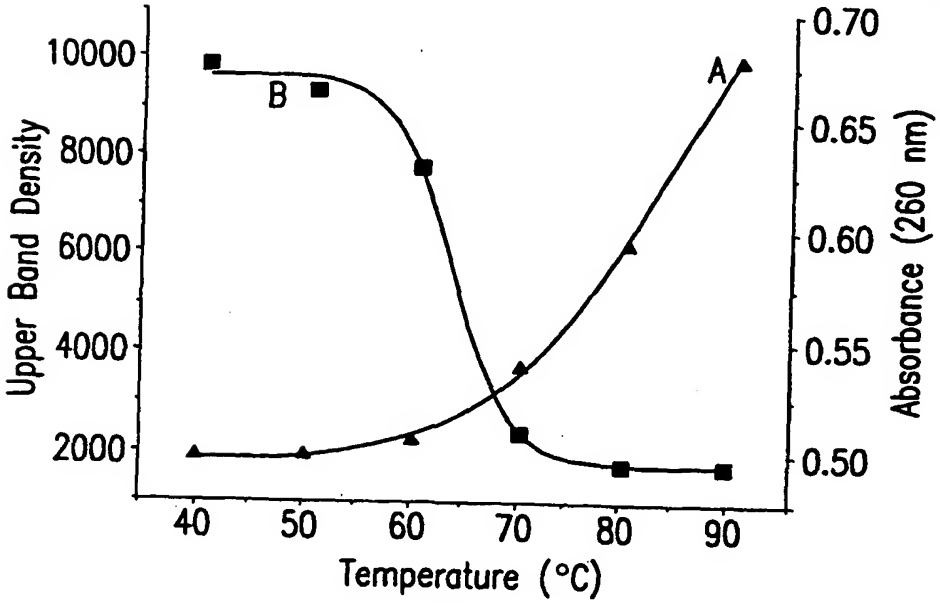
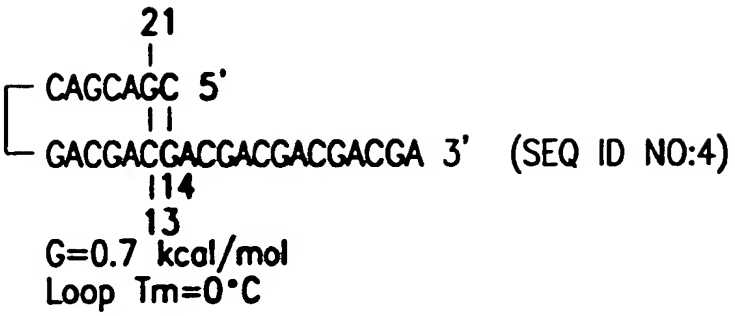

10000
8000
6000
4000
2000
Upper Band Density
B
A
0.70
0.65
0.60
0.55
0.50
Absorbance (260 nm)
40
50
60
70
80
90
Temperature (°C)

FIG.13

Stem Loop Structure:



G=0.7 kcal/mol
Loop Tm=0°C

Unstable Homodimer:

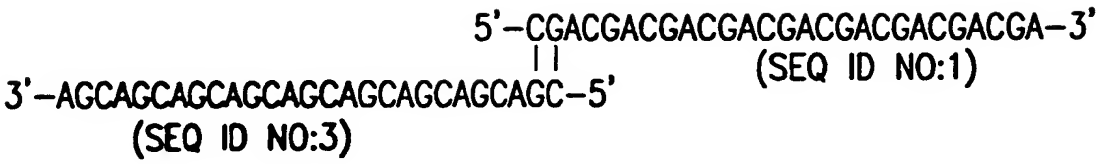


FIG.14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/03357

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) :C07H 21/00, 17/00; C12Q 1/68
US CL :536/24.3; 435/6
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/24.3; 435/6, 5, 91.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST, STN, MEDLINE, BIOSIS, CAPLUS, EMBASE, GENBANK
search terms: hybridization, sequencing, irradiation, photo-induction, nucleic

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98/12559 A1 (DEMERS) 26 March 1998, see entire document.	1-8
X	US 5,565,322 A (HELLER) 15 October 1996, see entire document.	1-8
---		---
Y		9-16
X	US 5,412,087 A (MCGALL et al) 02 May 1995, see entire document.	1-8
---		---
Y		9-16
Y	US 5,002,867 A (MACEVICZ) 26 March 1991, see entire document.	1-16
Y	US 5,795,722 A (LACROIX et al) 18 August, 1998, see entire document,	14-16

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 27 MARCH 2000	Date of mailing of the international search report 07 JUN 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer ARUN CHAKRABARTI Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)*

3NSDOCID: <WO___0047600A1_IA>

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/03357

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PEASE.A.C. et al. Light-generated Oligonucleotide Arrays for Rapid DNA Sequence Analysis. Proc. Natl. Acad.Sci. USA. 04 January 1994, Vol. 91. pages 5022-5026, see entire document.	1-16
X ---- Y	US 4,599,303 A (YABUSAKI et al) 08 July 1986, see entire document, especially column 5, line 8 to column 6, line 41.	9 ---- 10-16
X ----- Y	ZEHNDER et al. Cross-linking Hybridization Assay for Direct Detection of Factor V Leiden Mutation. Clinical Chemistry. September 1997. Vol. 43. No.9., pages 1703-1708, see entire document.	9-12 ------ 13-16

NSDOCID: <WO___0047600A1_IA>